A PROBLEM IN MINIMAX VARIANCE POLYNOMIAL EXTRAPOLATION

By A. Levine

University of Buenos Aires

1. Summary. For the problem of optimum prediction by means of kth degree polynomial regression, it is shown in [3] how to find the observation points and respective proportions of observations in the interval $[-1, 1]$ in order to obtain the minimax variance over the interval $[-1, t]$ of the predicted regression value for all $t \geqq t_1 > 1$; t_1 is the point outside the interval of observations at which the Chebyschev polynomial of degree k is equal to the maximum value of the variance of the least squares estimate in $[-1, 1]$. It is shown herein that if the observation points and proportions are chosen as specified in [3], then the maximum of the "least squares" variance in the interval $[-1, 1]$ is at -1. As a consequence, an equation is developed which permits the evaluation of t_1 as a function of k. Moreover, it is shown that $t_1 \to 1$ as $k \to \infty$, so that, for large k, the solution given in [3] yields an approximation to the minimax variance over the interval $[-1, t]$, all $t > 1$.

2. Introduction. Let $-1 \leqq x_i \leqq 1$, $i = 1, 2, \cdots, n$ denote the selected values of a variable x at which observations are to be made on a related variable y corresponding to those selected values. If $y(x_i)$ denotes the observed value of y corresponding to x_i, it will be assumed that the variables $y(x_i)$, $i = 1, 2, \cdots, n$, are uncorrelated random variables with common variance σ^2. It will also be assumed that the means of the y's lie on a polynomial curve of known degree k, that is, that

$$E[y(x_i)] = \beta_0 + \beta_1 x_i + \cdots + \beta_k x_i^k.$$

Let $\hat{y}(x)$ represent the least squares estimate of $E[y(x)]$ as a function of x and $V[\hat{y}(x)]$ the variance of the estimate. For the least squares estimator, which is also the minimum variance unbiased linear estimator, it is well known that the variance, $V[\hat{y}(x)]$, is equal to $x'(X'S^{-1}X)^{-1}x$ where $x' = (1, x, x^2, \cdots, x^k)$, S is the covariance matrix of the y's, and

$$X = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^k \\ 1 & x_2 & x_2^2 & \cdots & x_2^k \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^k \end{bmatrix}.$$

It is well-known (see [2] and [3]) that $V[\hat{y}(x)]$ will be minimized if the observations are concentrated at $k + 1$ points and that, consequently, $V[\hat{y}(x)]$ can be

Received 30 July 1965; revised 8 April 1966.

written in the form

$$V[\hat{y}(x)] = (\sigma^2/n)\textstyle\sum_{i=0}^{k} L_i^2(x)/p_i, \qquad p_i = n_i/n \tag{1}$$

where $L_i(x)$ is the Lagrange polynomial given by

$$L_i(x) = \textstyle\prod_{j\neq i}(x - x_j)/\prod_{j\neq i}(x_i - x_j),$$

and where n represents the total number of observations, while n_i represents the number of observations to be taken at x_i, $i = 0, 1, \cdots, k$.

The basic problem in minimax extrapolation at a point $t > 1$ is to find observation points and the corresponding proportions so that the resulting variance $V_t(x)$ has the property that its maximum in $[-1, t]$ is less than or equal to the maximum of the variance in this interval for any other choice of the set $\{x_i\}$ of observations and $\{p_i\}$ of proportions; i.e.,

$$\max_{x\varepsilon[-1,t]} V_t(x) = \min_{\{x_i\},\{p_i\}} \max_{x\varepsilon[-1,t]} V[\hat{y}(x)].$$

It has been shown in [3] that $V_t(x) = (\sigma^2/n)T_k(x; t)$ when $t \geqq t_1 > 1$; where t_1 is a point such that

$$\max_{-1\leqq x\leqq 1} T_k(x; t_1) = T_k(t_1; t_1)$$

and where

$$\begin{aligned} T_k(x; t) &= \textstyle\sum_{i=0}^{k} L_i^2(x)p_i^{-1}, \\ x_i &= -\cos(\pi i/k), \qquad i = 0, 1, \cdots, k \\ p_i &= |L_i(t)|/\textstyle\sum_{j=0}^{k}|L_j(t)|. \end{aligned} \tag{2}$$

The points $\{x_i\}$ given by the preceding formula are the points at which the kth degree Chebyschev polynomial takes on its maxima and minima in $[-1, 1]$ and are called the Chebyschev points. The next section develops an equation for the determination of t_1.

3. Determination of t_1.

LEMMA 1. *Let $x_i = -\cos(\pi i/k)$, $i = 0, 1, \cdots, k$, then each product $|\prod_{j\neq i}(x_i - x_j)|$ possesses the same value for $i = 1, 2, \cdots, k - 1$ and a value twice as large as this common value for $i = 0$ and $i = k$.*

The proof is carried out for $0 \leqq i \leqq k - 1$ by substituting the cosine representation of each x_j, x_i in the ratio

$$\left|\textstyle\prod_{j\neq i}(x_i - x_j)/\prod_{j\neq i+1}(x_{i+1} - x_j)\right|$$

and then using standard trigonometric identities to demonstrate the conclusion.

LEMMA 2. *If the Chebyschev points are chosen as observation points and if the Chebyschev weighting in (2) is used, then*

$$\max_{x\varepsilon[-1,1]} T_k(x; t) = T_k(-1; t), \qquad t > 1.$$

PROOF. From (2) it follows that when $t > 1$

$$p_0^{-1} = K[\prod_{i=1}^{k} (t - x_i)/|\prod_{i=1}^{k} (-1 - x_i)|]^{-1}$$

where $K = \sum_{i=0}^{k} |L_i(t)|$ and that

$$\begin{aligned} T_k(x; t) &= \sum_{i=0}^{k} L_i^2(x) p_i^{-1} \\ &= (\sum_{i=0}^{k} |L_i(t)|) \\ &\quad \cdot \{\sum_{j=0}^{k} \prod_{i\neq j} (x - x_i)^2/[|\prod_{i\neq j} (x_j - x_i)| \prod_{i\neq j} (t - x_i)]\}. \end{aligned}$$

Let $|\prod_{i\neq j} (x_j - x_i)| = a$ for $j = 1, 2, \cdots, k - 1$: Then, using Lemma 1,

$$\begin{aligned} T_k(x; t) &= (K/a)\{\sum_{j=1}^{k-1} \prod_{i\neq j} (x - x_i)^2/\prod_{i\neq j} (t - x_i) \\ &\quad + \prod_{i\neq k} (x - x_i)^2/2\prod_{i\neq k} (t - x_i) \\ &\quad + \prod_{i\neq 0} (x - x_i)^2/2\prod_{i\neq 0} (t - x_i)\} \\ &= (K/a)[\prod_{i=0}^{k} (t - x_i)]^{-1}\{\sum_{j=1}^{k-1} \prod_{i\neq j} (x - x_i)^2 (t - x_j) \\ &\quad + \tfrac{1}{2}[(t - 1)\prod_{i\neq k} (x - x_i)^2] + \tfrac{1}{2}[(t + 1)\prod_{i\neq 0} (x - x_i)^2]\} \\ &= (K/a)[\prod_{i=0}^{k} (t - x_i)]^{-1} A \end{aligned}$$

where A represents the coefficient of $(K/a)[\prod_{i=0}^{k} (t - x_i)]^{-1}$ in the preceding expression. Since $\max_{x\varepsilon[-1,1]} T_k(x; t) \geqq T_k(-1; t) = p_0^{-1}$, to prove the lemma it suffices to show that

$$p_0^{-1} \geqq \max_{[-1,1]} T_k(x; t) \qquad t > 1 \tag{3}$$

or, equivalently, since $|\prod_{i=1}^{k} (-1 - x_i)| = 2a$, that

$$2(t + 1)a^2 \geqq A. \tag{4}$$

To show (4), we note first that because $1 \geqq |x_j|$, $j = 0, 1, \cdots, k$, we have that $t + 1 \geqq t - x_j$ and so

$$\begin{aligned} A &\leqq (t + 1)\{\sum_{j=1}^{k-1} \prod_{i\neq j} (x - x_i)^2 + \tfrac{1}{2}\prod_{i\neq k} (x - x_i)^2 + \tfrac{1}{2}\prod_{i\neq 0} (x - x_i)^2\} \\ &= (t + 1)B; \qquad \text{all } t > 1. \end{aligned}$$

Hence it suffices to show that $2(t + 1)a^2 \geqq (t + 1)B$ or that

$$2a^2 \geqq B. \tag{5}$$

We shall demonstrate Inequality (5). Let $\varphi'(x)$ represent the derivative of the kth degree Chebyschev polynomial. Then (see [4])

$$\varphi'(x) = C\prod_{i=1}^{k-1} (x - x_i) \qquad C = 2^{k-1}k$$

$$x_i = -\cos(\pi i/k) \qquad i = 1, 2, \cdots, k - 1.$$

Using this expression to find the square of the derivative of $(x^2 - 1)\varphi'(x)$ in terms of B as well as in terms of both the product $\prod_{i\neq 0, i\neq k} (x - x_i)^2$ and the second derivative of $(x^2 - 1)\varphi'(x)$, we obtain after some calculations,

$$(6)\quad 2^{2(k-1)}k^2B = \{(d/dx)(x^2 - 1)\varphi'(x)\}^2 - (x + 1)^2\tfrac{1}{2}\varphi'^2(x) - (x - 1)^2 \tfrac{1}{2}(\varphi'^2(x)) - (x^2 - 1)\varphi'(x)(d^2/dx^2)[(x^2 - 1)\varphi'(x)] + 2(x^2 - 1)\varphi'^2(x).$$

We now make use of the well-known relation [4]

$$(7)\quad (x^2 - 1)\varphi''(x) + x\varphi'(x) - k^2\varphi(x) = 0.$$

Using (7) we have that

$$(8)\quad \begin{aligned}(d/dx)(x^2 - 1)\varphi'(x) &= 2x\varphi'(x) + (x^2 - 1)\varphi''(x)\\ &= x\varphi'(x) + k^2\varphi(x).\end{aligned}$$

Hence

$$(9)\quad (d^2/dx^2)(x^2 - 1)\varphi'(x) = x\varphi''(x) + (k^2 + 1)\varphi'(x).$$

Putting (8), (9) in (6) we have

$$(10)\quad \begin{aligned}2^{2(k-1)}k^2B = D = \{x\varphi'(x) &+ k^2\varphi(x)\}^2 - \tfrac{1}{2}[(x + 1)^2\varphi'^2(x)]\\ &- \tfrac{1}{2}(x - 1)^2\varphi'^2(x) - x(x^2 - 1)\varphi''(x)\varphi'(x)\\ &- (x^2 - 1)\varphi'^2(x)(k^2 + 1) + 2(x^2 - 1)\varphi'^2(x).\end{aligned}$$

Using (7) we can substitute the value of $(x^2 - 1)\varphi''(x)$ in (10) to obtain

$$\begin{aligned}D &= x^2\varphi'^2(x) + 2k^2x\varphi(x)\varphi'(x) + k^4\varphi^2(x) - x^2\varphi'^2(x) - k^2x\varphi(x)\varphi'(x)\\ &\quad - x^2\varphi'^2(x) - \varphi'^2(x) - (x^2 - 1)\varphi'^2(x)(k^2 + 1) + 2(x^2 - 1)\varphi'^2(x)\\ &= k^4\varphi^2(x) + k^2x\varphi'(x)\varphi(x) - k^2(x^2 - 1)\varphi'^2(x) + 2(x^2 - 1)\varphi'^2(x)\\ &\quad - 2x^2\varphi'^2(x).\end{aligned}$$

Thus, we have

$$D = k^4\varphi^2(x) + k^2x\varphi(x)\varphi'(x) - k^2(x^2 - 1)\varphi'^2(x) - 2\varphi'^2(x).$$

From [4] we know that if $x = \cos\theta$ then

$$\begin{aligned}\varphi(x) &= \cos k\theta\\ \varphi'(x) &= k\sin k\theta/\sin\theta \qquad\qquad \pi \geqq \theta \geqq 0.\end{aligned}$$

Then

$$D = k^4\cos^2 k\theta + k^3\cos\theta[(\sin k\theta\cos k\theta)/\sin\theta] + k^4\sin^2 k\theta - 2k^2\sin^2 k\theta/\sin^2\theta$$

or

$$D \leqq k^4 + k^3\cos\theta(\sin k\theta\cos k\theta/\sin\theta).$$

Note that the maximum of $\cos\theta$, $\cos k\theta$ is equal to 1 and that the maximum of

$\sin k\theta/\sin\theta$ is k when $\theta \varepsilon [0, \pi]$. Hence $D \leqq 2k^4$ or $B \leqq 2k^2/2^{2(k-1)}$. To complete the proof we must show that

$$(11) \qquad 2k^2/2^{2(k-1)} \leqq 2a^2.$$

We note that $a = \prod_{j \neq i} |x_i - x_j|$, $i \neq 0$, $i \neq k$. Thus

$$a = (x_i^2 - 1)\varphi''(x_j)/C; \qquad i \neq 0, i \neq k.$$

From (7)

$$Ca = |-x_i\varphi'(x_i) + k^2\varphi(x_i)|$$

But $\varphi'(x_i) = 0$, $\varphi(x_i) = +1$ or -1. Thus $C^2a^2 = k^4$; but $C = 2^{k-1}k$. Hence $a = k/2^{k-1}$. Substituting back into (11) we have $2k^2/2^{2(k-1)} \leqq 2k^2/2^{2(k-1)}$ which completes the proof.

Theorem 1. *The value of* t_1 *can be found by solving the equation*

$$(12) \quad \sum_{i=0}^{k} |L_i(t_1)| = |\prod_{i=1}^{k} (-1 - x_i)|/|\prod_{i=1}^{k} (t_1 - x_i)| \quad x_i = -\cos(\pi i/k).$$

Proof. This equation is obtained from Lemma 2 by applying it to the equation defining t_1 just before Formula (2). It should be remembered that $T_k(-1; t) = p_0^{-1}$ when $t = t_1$.

4. Determination of the range of t_1.

Theorem 2. *The solution of Equation* (12) *above is in the interval* (1, 2].

Proof. Reference [3] has shown that t_1, the solution to (12), is greater than 1. To demonstrate that $t_1 \leqq 2$ we note once more that

$$(13) \qquad T_k(x; t) = \sum_{i=0}^{k} L_i^2(x)p_i^{-1} = \sum_{i=0}^{k} [\prod_{j\neq i} (x - x_j)^2/D_i]$$

where D_i is a positive constant depending only on i and t. Suppose first that $k > 1$. By symmetry of the "Chebyschev points", $x_i = -\cos(\pi i/k)$, we have that

$$T_k(x; t) = \sum_{i=0}^{k} [(x - x_j)^2 \prod_{j\neq i}^{0\leqq j\leqq (k+2)/2} (x^2 - x_j^2)^2/D_i], \qquad k > 1, k \text{ even}$$

$$T_k(x; t) = \sum_{i=0}^{k} [(x - x_i)^2 \prod_{j\neq i}^{0\leqq j\leqq (k-1)/2} (x^2 - x_j^2)^2/D_i], \qquad k > 1, k \text{ odd.}$$

Since $x_j^2 \varepsilon [0, 1]$, if $x \varepsilon [-1, 1]$ then $x^2 \varepsilon [0, 1]$ and $\prod_{j\neq i}^{0\leqq j\leqq k/2} (x^2 - x_j^2)^2 < 1$, $(x + x_i)^2 \leqq 4$. But if $t > 2$, then $\prod_{j\neq i}^{0\leqq j\leqq k/2} (t^2 - x_j^2)^2 > (9)^{k/2} > 4$ and $(t - x_i)^2 \geqq 1$. Hence, terms of $T_k(t; t)$, $t > 2$, are larger or equal to the terms, term by term, of $T_k(x; t)$. Consequently, $T_k(t; t) \geqq T_k(x; t)$ when $t > 2$: thus t_1 cannot be greater than 2.

We now prove that there exists a solution to (12) in (1, 2]. First, let $k > 1$. When $t_1 = 1$, the left side of (12) is equal to 1 while the right side is unbounded. We have already shown that if $t_1 > 2$, then

$$T_k(t_1; t_1) > p_0^{-1} \qquad k > 1$$

or that the left side of (12) is greater than the right side. Since both sides are continuous in t_1, their graphs must intersect in (1, 2] thus proving that the solution is in this interval. To prove the above result for $k = 1$, (12) can easily be solved directly yielding $t_1 = 2$.

It follows immediately from the above results that $\lim_{k\to\infty} t_1 = 1$. Moreover, it is a simple manner to demonstrate that the convergence is at most of order k^{-1}, for (12) can be written in the form

$$\varphi(t_1) = 2ak2^{k-1}/(t_1 - 1)\varphi'(t_1). \tag{14}$$

Hence, for t_1 close to 1 ($t_1 - 1 = \epsilon > 0$),

$$\varphi(t_1) = \varphi(1) + \epsilon\varphi'(1) + o(\epsilon). \tag{15}$$

Using (7) and noting that $\varphi(1) = 1$, we have

$$\varphi(t_1) = 1 + \epsilon k^2 + o(\epsilon).$$

Clearly $\varphi'(t_1)$ is an increasing function of t_1 when $t_1 \geqq 1$. Hence from (14), (15),

$$\varphi(t_1) = 1 + \epsilon k^2 + o(\epsilon) = (2ak2^{k-1}/\epsilon\varphi'(t_1)) + o(\epsilon) \leqq (2ak2^{k-1}/\epsilon\varphi'(1)) + o(\epsilon)$$
$$= (2a2^{k-1}/\epsilon k) + o(\epsilon).$$

Since $a = (k/2^{k-1})$ we have $1 + \epsilon k^2 \leqq (2/\epsilon) + |o(\epsilon)|$. Consequently, $0 < \epsilon \leqq k^{-2}\{\frac{1}{2}[(1 + [8 - 4\,|o(\epsilon)|]k^2)^{\frac{1}{2}} - 1]\}$, so that for all k, $0 < \epsilon \leqq (2 - |o(\epsilon)|)^{\frac{1}{2}}/k$, which demonstrates that ϵ is at most of order k^{-1}.

The above remarks show that for large k, the solution given in [3] becomes a good approximation to the minimax variance in the interval $(-1, t]$ all $t > 1$. The convergence of t_1 to 1 is rather rapid even for small k as is indicated below by the partial reproduction of one of the tables in [1]:

k	1	2	5	10	100
t_1	2	1.44061	1.13185	1.04918	1.00133

Finally, we observe that the solution t_1 to (12) is unique. This conclusion i obtained at once from (14) by using the fact that both $\varphi(t)$ and $\varphi'(t)$ are increasing functions of t for $t > 1$.

REFERENCES

[1] Alfonso, F. J. (1965). Tablas para diseño optimo minimax en predicción polinomial. Inst. de Cálculo, Universidad de Buenos Aires, Publicación No. 9.

[2] de la Garza, A. (1954). Spacing of information in polynomial regression. *Ann. Math. Statist.* **25** 123–130.

[3] Hoel, P. G. and Levine, A. (1964). Optimal spacing and weighting in polynomial prediction. *Ann. Math. Statist.* **35** 1553–1560.

[4] *Tables of Chebyschev Polynomials*, National Bureau of Standards, Applied Mathematics Series, No. 9.