The Annals of Statistics
1980, Vol. 8, No. 5, 1132–1141

D_S-OPTIMAL DESIGNS FOR POLYNOMIAL REGRESSION USING CONTINUED FRACTIONS[1]

BY W. J. STUDDEN

Purdue University

Consider a polynomial regression of degree n on an interval. Explicit optimal designs are given for minimizing the determinant of the covariance matrix of the least squares estimators of the highest s coefficients. The designs are calculated using continued fractions.

1. Introduction. Consider a polynomial regression situation on $[-1, 1]$. For each x or "level" in $[-1, 1]$ an experiment can be performed whose outcome is a random variable $Y(x)$ with mean value $\sum_{i=0}^{n} \beta_i x^i$ and variance σ^2, independent of x. The parameters β_i, $i = 0, 1, \cdots, n$ and σ^2 are unknown. An experimental design is a probability measure ξ on $[-1, 1]$. If N observations are to be taken and ξ concentrates mass $\xi(i)$ at the points x_i, $i = 1, 2, \cdots, v$ and $\xi(i)N = n_i$ are integers, the experimenter takes N uncorrelated observations, n_i at each x_i, $i = 1, 2, \cdots, v$. The covariance matrix of the least squares estimates of the parameters β_i is then given by $(\sigma^2/N)M^{-1}(\xi)$ where $M(\xi)$ is the information matrix of the design with elements $m_{ij} = \int_{-1}^{1} x^{i+j}\, d\xi(x)$. For an arbitrary probability measure or design some approximation would be needed in applications.

Let $f'(x) = (1, x, x^2, \cdots, x^n)$ and $d(x, \xi) = f'(x)M^{-1}(\xi)f(x)$ when $M(\xi)$ is nonsingular. It is known for general regression functions, see Kiefer and Wolfowitz (1960), that the design minimizing $\sup_x d(x, \xi)$ and the design maximizing the determinant $|M(\xi)|$ are the same. This is referred to as the D-optimal design. In the polynomial case it concentrates equal mass $(n + 1)^{-1}$ on each of the $n + 1$ zeros of $(1 - x^2)P_n'(x) = 0$, where P_n is the nth Legendre polynomial, orthogonal to the uniform measure on $[-1, 1]$. The solution of the separate problems for polynomial regression was discovered earlier by Hoel (1958) and Guest (1958) leading Kiefer and Wolfowitz to their equivalence theorem.

It is also known (see Kiefer and Wolfowitz (1959)) that the design that minimizes the variance of the highest coefficient β_n concentrates mass proportional to $1 : 2 : 2 : \cdots :2 : 1$ (nearly equal) on the zeros of $(1 - x^2)T_n'(x) = 0$ where T_n is the Chebyshev polynomial of the first kind. These are orthogonal with respect to $(1 - x^2)^{-\frac{1}{2}}$.

The purpose of this paper is to consider the D_s-optimal design which minimizes the determinant of the covariance matrix of the least squares estimates of the highest s parameters $\beta_{r+1}, \beta_{r+2}, \cdots, \beta_n$, where $n - r = s$. The estimation of all

Received January 1979; revised May 1979.

[1]This research was supported by NSF Grant No. MCS76-08235 A01.

AMS 1970 *subject classification*. Primary 62K05, 62J05.

Key words and phrases. D_s-optimal, polynomial regression, continued fractions.

coefficients is the D_{n+1} or D-optimal situation. It will be easily seen below that the D_n-optimal and D-optimal designs are the same.

Let $f'(x) = (f_1'(x), f'_2(x))$ where $f_1'(x) = (1, x, \cdots, x^r)$ and $f_2'(x) = (x^{r+1}, \cdots, x^n)$ and let the information matrix $M(\xi)$ have a similar decomposition

$$M(\xi) = \begin{vmatrix} M_{11}(\xi) & M_{12}(\xi) \\ M_{21}(\xi) & M_{22}(\xi) \end{vmatrix}. \tag{1.1}$$

The covariance matrix of the estimates for $\beta_{r+1}, \cdots, \beta_n$ is proportional to the inverse of

$$\Sigma = \Sigma(\xi) = M_{22}(\xi) - M_{21}(\xi)M_{11}^{-1}(\xi)M_{12}(\xi). \tag{1.2}$$

The problem is to maximize the determinant of $\Sigma(\xi)$. Corresponding to the ordinary D-optimal situation the design maximizing $|\Sigma(\xi)|$ also minimizes the supremum over $[-1, 1]$ of

$$d_s(x, \xi) = (f_2(x) - A(\xi)f_1(x))'\Sigma^{-1}(\xi)(f_2(x) - A(\xi)f_1(x)) \tag{1.3}$$

where $A(\xi) = M_{21}M_{11}^{-1}$. Moreover for the optimal design ξ_s

$$d_s(x, \xi_s) \leqslant s. \tag{1.4}$$

To find the maximum of $|\Sigma(\xi)|$ we use the result that $|\Sigma(\xi)| = |M(\xi)||M_{11}(\xi)|^{-1}$. Note that $r = n - s = 0$ corresponds to the D-optimal case since $M_{11} = 1$. The quantity $d_n(x, \xi)$ in (1.4) is not, however, equal to $d(x, \xi)$ defined above. In fact, as mentioned at the beginning of Section 5, equation (4.5) shows that

$$d_n(x, \xi) = d(x, \xi) - 1.$$

In the following the moments m_{ij} and the determinants $|M(\xi)|$ and $|M_{11}(\xi)|$ will be expressed in an appropriate form using certain "canonical moments." The maximization of the determinant $|\Sigma(\xi)|$ then becomes very easy. The solution is then converted back to the moments $m_{ij} = \int x^{i+j}\, d\xi(x)$ and the design ξ_s. Section 2 contains the maximization of the determinant $|\Sigma(\xi)|$. The relationship between the ordinary moments and the canonical moments is described in Section 3. This relationship involves some simple recursive formulas which also relate the "generating function" for the ordinary moments with its continued fraction expansion. The continued fractions are used more fully in Section 4 in obtaining the support of the D_s-optimal design. Some examples are given in each of the sections and in Section 5.

The problem considered here is described for polynomial regression on $[-1, 1]$. It can readily be seen that it is invariant under a simple linear transformation onto any interval $[a, b]$. In the sections below it will be seen that certain expressions are more readily available and possibly simpler for the interval $[0, 1]$. We have chosen the interval $[-1, 1]$ because the classical orthogonal polynomials are usually given on this interval and the details of some of our examples are easier to handle because of symmetry considerations.

The results in Theorems 2.1, 4.1 and 4.2 provide some simple and useful D_s-optimal designs. A comparison of these with others in the literature for polynomial regression is being considered.

2. Maximization of $|\Sigma(\xi)|$. The maximization of $|\Sigma(\xi)|$ defined in (1.2) is done in terms of simple expressions for $|M(\xi)|$ using canonical moments. For an arbitrary design or probability measure ξ on $[-1, 1]$ let

$$c_k = \int_{-1}^{1} x^k \, d\xi(x), \qquad k = 0, 1, 2, \cdots.$$

For a given set of moments $c_0, c_1, \cdots, c_{i-1}$ let c_i^+ denote the maximum of the ith moment $\int x^i \, d\mu(x)$ over the set of all measures μ having moments $c_0, c_1, \cdots c_{i-1}$. Similarly let c_i^- denote the corresponding minimum. The canonical moments are defined by

$$p_i = \frac{c_i - c_i^-}{c_i^+ - c_i^-} \qquad i = 1, 2, \cdots. \tag{2.1}$$

Note that $0 \leqslant p_i \leqslant 1$. The canonical moments for the "Jacobi" measure $(1 + x)^\alpha (1 - x)^\beta$, along with other considerations, are given in Skibinsky (1969). For $\alpha = \beta$ it is shown that $p_i = \frac{1}{2}$ for i odd and

$$p_i = \tfrac{1}{2}\left(\frac{i}{i + 2\alpha + 1}\right)$$

for i even. Note that the usual arc-sin law or the measure corresponding to the ordinary Chebyshev polynomials has $\alpha = -\frac{1}{2}$ and the canonical moments are $p_i \equiv \frac{1}{2}$.

THEOREM 2.1. *The determinant $|\Sigma(\xi)|$ in* (1.2) *is maximized by the unique design ξ whose canonical moments are given by*

$$\begin{aligned}
p_i &= \tfrac{1}{2} && \text{for } i \text{ odd;} \\
p_{2i} &= \tfrac{1}{2} && i = 1, 2, \cdots, r; \\
&= \frac{n - i + 1}{2n - 2i + 1} && i = r + 1, \cdots, n - 1; \\
&= 1 && i = n.
\end{aligned} \tag{2.2}$$

PROOF. Let Δ_{2m} denote the determinant

$$\Delta_{2m} = \begin{vmatrix} c_0 & c_1 & \cdots & c_m \\ c_1 & c_2 & \cdots & c_{m+1} \\ \vdots & \vdots & & \vdots \\ c_m & c_{m+1} & \cdots & c_{2m} \end{vmatrix} \qquad m = 0, 1, 2, \cdots.$$

The maximizing canonical moments can easily be found once we express the determinant Δ_{2m} in terms of the canonical moments. These are found using Skibinsky (1968). First consider transforming the measure ξ on $X = [-1, 1]$ to

$Z = [0, 1]$ using a simple linear transformation $x = 2z - 1$ and let the corresponding moments on $Z = [0, 1]$ be denoted by b_i. It is shown in Skibinsky (1969) that the canonical moments are invariant and the determinant Δ_{2m} for the moments b_i is given by

(2.3) $$\Pi_{i=1}^{m}(\eta_{2i-1}\eta_{2i})^{m+1-i}$$

where

(2.4) $$\eta_0 = q_0, \eta_j = q_{j-1}p_j \qquad j = 1, 2, \cdots$$

and $p_j + q_j = 1$ for all j. The determinant Δ_{2m} for the moments $c_i = \int_{-1}^{1} x^i \, d\xi(x) = \int_0^1 (2z - 1)^i \, d\xi(2z - 1)$ is then a power of two times the quantity (2.3). Therefore, the maximization of the determinant $|\Sigma(\xi)|$ is equivalent to maximizing

$$\frac{\Pi_{i=1}^{n}(\eta_{2i-1}\eta_{2i})^{n+1-i}}{\Pi_{i=1}^{r}(\eta_{2i-1}\eta_{2i})^{r+1-i}}.$$

Some fairly simple algebra shows the answer to be given by (2.2).

The uniqueness is a consequence of the fact that the values $p_i, i = 0, \cdots, 2n$ uniquely determine $c_0, c_1, \cdots, c_{2n}$ and the value $p_{2n} = 1$ implies that $(c_0, c_1, \cdots, c_{2n})$ is a boundary point of the corresponding moment space for which the measure ξ is unique.

3. Ordinary and canonical moments. The relationship between the two types of moments is expressed by the use of certain simple recursive relationships which relate the power series or generating function for the moments to a continued fraction expansion.

For a probability measure μ on [0, 1] with moments b_i the power series

(3.1) $$P(w) = \Sigma_{i=0}^{\infty} b_i w^i$$

has a continued fraction expansion of the form

(3.2) $$\frac{1}{1} - \frac{\eta_1 w}{1} - \frac{\eta_2 w}{1} - \cdots$$

where the quantities η_i are given in (2.4). More details of these considerations are given in Seall and Wetzel (1959) or Wall (1940). All of the series and continued fractions we consider will be convergent. These questions will not actually concern us since our interest will be in finite sections of the expansions and the formal relationships between the coefficients.

Define the numbers S_{ij} recursively by $S_{0j} = 1, j = 0, 2, \cdots$ and for $i \leqslant j$

(3.3) $$S_{ij} = \Sigma_{k=i}^{j} \eta_{k-i+1} S_{i-1,k} \qquad i,j = 1, 2, \cdots.$$

The corresponding moments b_i are given by

(3.4) $$b_m = S_{mm} \qquad m = 1, 2, \cdots.$$

These are taken from Skibinsky [1969]. The moments c_m of the translated measure on $[-1, 1]$ can be obtained from the b_i by the relation $c_i = \int_0^1 (2z - 1)^i \, d\mu(z)$.

The moments c_m can be found from the η_i more directly using (3.3) for the canonical moments of interest given in (2.2) where the odd values are all $p_i = \frac{1}{2}$. The power series

$$P(w) = \Sigma_{i=0}^{\infty} c_i w^i$$

has a continued fraction expansion of the form (see Seall and Wetzel (1959))

$$\frac{1}{e_1 w + 1} - \frac{d_1 w^2}{e_2 w + 1} - \frac{d_2 w^2}{e_3 w + 1} - \cdots$$

where

(3.5)
$$d_i = 4m_i(1 - m_{i-1})l_i(1 - l_i) \qquad i = 1, 2, \cdots$$
$$e_i = 1 - 2m_{i-1}(1 - l_{i-1}) - 2(1 - m_{i-1})l_i \qquad i = 1, 2, \cdots$$
$$m_i = p_{2i} \text{ and } l_i = p_{2i-1} \qquad i = 1, 2, \cdots.$$

For the p_i given in (2.2) the odd $p_{2i-1} = \frac{1}{2}$ so that $e_i = 0$ and $d_i = q_{2i-1}p_{2i} = \eta_{2i}$. Thus the recursive relations (3.3) can be used to calculate the even moments c_{2i}, the odd moments being zero.

Consider for example the uniform probability measure on $[-1, 1]$ or on $[0, 1]$. We noted previously that for this measure $p_{2i+1} = \frac{1}{2}$ and $p_{2i} = i/(2i + 1)$. Therefore we have

$$p_1 = \tfrac{1}{2}, p_2 = \tfrac{1}{3}, p_3 = \tfrac{1}{2}, p_4 = \tfrac{2}{5}, p_5 = \tfrac{1}{2}, p_6 = \tfrac{3}{7}.$$

The first few values for S_{ij} in (3.3) give

$$b_1 = p_1$$
$$b_2 = p_1(p_1 + q_1 p_2)$$
$$b_3 = p_1[p_1(p_1 + q_1 p_2) + q_1 p_2(p_1 + q_1 p_2 + q_2 p_3)].$$

Substituting the values in gives $b_1 = \frac{1}{2}$, $b_2 = \frac{1}{3}$, $b_3 = \frac{4}{4}$. On the interval $[-1, 1]$ the same three formulas give the six moments $c_0, c_1, \cdots, c_6$ using $c_{2i+1} = 0$ and

$$c_2 = p_2$$
$$c_4 = p_2(p_2 + q_2 p_4)$$
$$c_6 = p_2[p_2(p_2 + q_2 p_4) + q_2 p_4(p_2 + q_2 p_4 + q_4 p_6)].$$

This gives $c_2 = \frac{1}{3}$, $c_4 = \frac{1}{5}$, $c_6 = \frac{1}{7}$.

4. The moments c_j and the design ξ_s.

In this section we prove the following two theorems.

THEOREM 4.1. *The support of the D_s-optimal design ξ_s consists of the points ± 1 and the $n - 1$ zeros of*

(4.1)
$$\rho_s'(x)\tau_{r+1}'(x) - \alpha_s \rho_{s-1}'(x)\tau_r'(x) = 0$$

where $r + s = n$,

(4.2)
$$\alpha_s = \frac{1}{2}\frac{s-1}{2s-1} \qquad s = 1, 2, \cdots, n$$

and ρ_i' and τ_i' have leading coefficients one and are proportional to the derivatives of the Legendre and Chebyshev polynomials $P_i(x)$ and $T_i(x)$.

THEOREM 4.2. *The weights of the D_s-optimal design attached to the points $x_0, x_1, \cdots, x_n$ of* (4.1) *are given by*

$$\xi_s(i) = \frac{2}{2n + 1 + U_{2r}(x_i)} \quad i = 0, 1, \cdots, n, r + s = n \tag{4.3}$$

where $U_{2r}(x)$ is the Chebyshev polynomial of the second kind $U_k(x) = (\sin(k+1)\theta/\sin\theta)$, $x = \cos\theta$.

The support points of the D_s-optimal design in the polynomial in (4.1) can be expressed in another form. If $Q_k(x), k = 0, 1, \cdots$ are the polynomials of the indicated degree orthogonal to $(1 - x^2)\,d\xi_s(x)$ then the support points are the zeros of $(1 - x^2)Q_{n-1}(x) = 0$. See Karlin and Studden (1966), Chapter 4. The polynomial $Q_{n-1}(x)$ can be obtained by calculating the moments c_i from the p_i as explained in Section 3. The required polynomial can be expressed as

$$Q_{n-1}(x) = \begin{vmatrix} c_0 - c_2 & c_1 - c_3 & \cdots & c_{n-2} - c_n & 1 \\ c_1 - c_3 & c_2 - c_4 & & c_{n-1} - c_{n+1} & x \\ \vdots & & & & \\ c_{n-1} - c_{n+1} & \cdots & & c_{2n-3} - c_{2n-1} & x^{n-1} \end{vmatrix} \tag{4.4}$$

which must be proportional to the polynomial in (4.1). Our proof of Theorem 4.1 goes directly from the p_i value to (4.1) using continued fractions.

PROOF OF THEOREM 4.2. The weights $\xi_s(i)$ are obtained by using the fact that for each point x_i in the support of the optimal design, equality must hold in (1.4). That is $d_s(x_i, \xi_s) = s$. The quantity $d_s(x, \xi)$ can be rewritten in the form

$$d_s(x, \xi) = f'(x)M^{-1}(\xi)f(x) - f_1'(x)M_{11}^{-1}(\xi)f_1(x). \tag{4.5}$$

We now change the basis for the two terms on the right-hand side. For the first term we use the Lagrange polynomials $l_j(x), j = 0, 1, \cdots, n$ associated with the optimal points x_i. The $l_j(x)$ are the polynomials of degree n satisfying $l_j(x_i) = \delta_{ij}$ $i, j = 0, 1, \cdots, n$. If ξ has mass $\xi(i)$ on x_i then

$$f'(x)M^{-1}(\xi)f(x) = \Sigma_{i=0}^{n} \frac{l_i^2(x)}{\xi(i)}.$$

For the second term, note that the canonical moments up to order $2r$ (those used in M_{11}) are $p_i = \frac{1}{2}$. These correspond to the Chebyshev measure $(1 - x^2)^\alpha$ with $\alpha = -\frac{1}{2}$. For the second term use the polynomials, $1, 2^{1/2}T_1(x), 2^{1/2}T_2(x), \cdots$ which are the orthonormal Chebyshev polynomials. In this case

$$f_1'(x)M_{11}^{-1}(\xi_s)f_1(x) = 1 + 2\Sigma_{j=1}^{r} T_j^2(x).$$

Using the fact that $d_s(x_i, \xi_s) = s$ we then find that

$$s = \frac{1}{\xi(i)} - \left(1 + 2\Sigma_{j=1}^{r} T_j^2(x_i)\right).$$

The solution $\xi_s(i)$ in (4.3) is then obtained using the fact that $T_j(x) = \cos j\theta$ where $x = \cos\theta$ and (see Jolley (1961)]

$$\begin{aligned}\Sigma_{j=1}^{r} T_j^2(x) &= \tfrac{1}{2}\left(r + \frac{\cos(r+1)\theta \sin r\theta}{\sin\theta}\right) \\ &= \tfrac{1}{2}\left(r - \tfrac{1}{2} + \frac{U_{2r}(x)}{2}\right).\end{aligned}$$

Proof of Theorem 4.1. For the optimal design ξ_s the canonical moment $p_{2n} = 1$ implies that c_{2n} has a maximum value given the set $c_0, c_1, \cdots, c_{2n-1}$. The support points of ξ_s are then $x = \pm 1$ and the zeros of the polynomial $Q_{n-1}(x)$ orthogonal to the measure $(1 - x^2)\, d\xi_s(x)$. See Karlin and Studden (1966), Ch. IV. These polynomials will be obtained using continued fractions.

Consider a given set of moments $c_0, c_1, \cdots$ corresponding to some ξ and take

$$\Sigma_{i=0}^{\infty} c_i w^{-i-1} = \frac{c_0}{w} + \frac{c_1}{w^2} + \frac{c_2}{w^3} + \cdots$$

in a continued fraction form

$$\frac{1}{A_1 w + B_1} - \frac{C_2}{A_2 w + B_2} - \cdots - \frac{C_k}{A_k w + B_k} - \cdot \tag{4.6}$$

The polynomial orthogonal to ξ of degree k is given by the denominator of the kth convergent. That is, take the expansion (4.6) only up to the kth term and express it as a ratio of two polynomials. The denominator is the required polynomial. See Szego (1959), page 55.

For the optimal design ξ_s we have $c_{2i-1} = 0$ and $\Sigma c_{2i} w^{2i}$ has the expansion

$$\frac{1}{1} - \frac{d_1 w^2}{1} - \frac{d_2 w^2}{1} - \cdots \tag{4.7}$$

where $d_i = m_i(1 - m_{i-1}), i = 1, 2, \cdots, m_0 = 0$ and $m_i = p_{2i}$. The moments for the measure $(1 - x^2)\, d\xi_s(x)$ are $c_i - c_{i+2}$. Using Wall (1940) the expansion for $\Sigma(c_{2i} - c_{2i+2})w^{2i}$ can be obtained from (4.7) and is

$$\frac{c_0 - c_2}{1} - \frac{\alpha_1 w^2}{1} - \frac{\alpha_2 w^2}{1} - \cdots \tag{4.8}$$

where $\alpha_i = m_i(1 - m_{i+1}), i = 1, 2, \cdots$ and $m_i = p_{2i}$. We require the polynomials in the denominators of the convergents of the continued fraction expansion of

$$\Sigma_{i=0}^{\infty} \frac{c_{2i} - c_{2i+2}}{w^{2i+1}}. \tag{4.9}$$

Replacing w by $1/w$ in (4.8), multiplying by $1/w$ and making some "equivalence transformations" (see Wall (1948)), we can express the expansion for (4.9) as

$$\frac{c_0 - c_2}{w} - \frac{\alpha_1}{w} - \frac{\alpha_2}{w} - \cdots - \frac{\alpha_k}{w} - \cdots. \tag{4.10}$$

The proof now involves terminating the expansion (4.10) at $k = n - 2$ and finding the resulting denominator. Basically the α_i values break into two parts. Those in the first part are associated with the Chebyshev polynomials or arc-sin law and those in the second half are associated with Lebesgue measure. Certain formulas in continued fractions allow us to write the resulting polynomial in the form (4.1).

A number of facts are required concerning the quantities $\alpha_i = p_{2i}(1 - p_{2i+2})$. For the Chebyshev or arc-sin measure proportional to $(1 - x^2)^\alpha$ with $\alpha = -\frac{1}{2}$, the canonical moments are all $p_i = \frac{1}{2}$ (See Skibinsky (1969)). These are the same as the first part of the canonical moments corresponding to the optimal ξ_s. Thus the values $\alpha_1, \alpha_2, \cdots, \alpha_{r-1}$ are the same values that would have been obtained by starting in (4.7) with the Chebyshev measure. The polynomials for the difference moments as in (4.9) correspond, at least for the classical polynomials, to the derivative of the corresponding polynomial for the given moments. Thus, if we truncate the expansion (4.10) at $k \leqslant r - 1$ the corresponding polynomial is $T'_{k+2}(w)$ with leading coefficient equal to one.

The value α_r is given by (4.2). The remaining set of values $\alpha_{r+1}, \alpha_{r+2}, \cdots, \alpha_{n-2}$ are related to the uniform measure which has $p_{2i} = i/(2i + 1)$. Let the corresponding α_i obtained as above for the uniform measure be denoted by α'_i. It can then be checked that the reversed values $\alpha_{n-2}, \alpha_{n-3}, \cdots, \alpha_{r+1}$ are $\alpha'_1, \alpha'_2, \cdots$. Thus $\alpha'_i = \alpha_{n-1-i}, i = 1, 2, \cdots, n - r - 2$.

The proof can now be completed by taking the terms in (4.10) up to $k = n - 2$, and applying certain basic formulas in continued fractions. These are given in Perron (1954) and are as follows. Let

$$b_0 + \frac{a_1}{b_1} + \cdots + \frac{a_v}{b_v} = \frac{A_v}{B_v}$$

and define

$$B_v = K\begin{pmatrix} a_2 a_3 \cdots a_v \\ b_1 b_2 \cdots b_v \end{pmatrix}.$$

Then

$$K\begin{pmatrix} a_2 a_3 \cdots a_v \\ b_1 b_2 \cdots b_v \end{pmatrix} = K\begin{pmatrix} a_v a_{v-1} \cdots a_2 \\ b_v b_{v-1} \cdots b_1 \end{pmatrix}. \tag{4.11}$$

Moreover, if

$$B_{v\lambda} = K\begin{pmatrix} a_{\lambda+2} a_{\lambda+3} \cdots a_{\lambda+v} \\ b_{\lambda+1} b_{\lambda+2} \cdots b_{\lambda+v} \end{pmatrix}$$

then

$$B_{v+\lambda-1} = B_{\lambda-1} B_{v,\lambda-1} + a_\lambda B_{\lambda-2} B_{v-1,\lambda}. \tag{4.12}$$

The polynomial in (4.1) is obtained from (4.12) if we let $\lambda = r + 1, v = -r + 1, a_{i+1} = -\alpha_i, b_i = w$ and use (4.11) on the two B terms with double subscripts.

5. Examples. From Theorems 4.1 and 4.2 the D_n-optimal design has equal weight $1/(n+1)$ on the zeros of $(1-x^2)P_n'(x) = 0$. It was mentioned earlier that the D_n and D-optimal are the same. That is, the design maximizing $|M|/|M_{11}|$ and $|M|$ are the same since $f_0 = 1$ and $M_{11} = 1$. The D-optimal design has the property that $d(x, \xi) = f'(x)M^{-1}(\xi)f(x) \leqslant n + 1$ while by (1.4) $d_n(x, \xi) \leqslant n$ for the D_n-optimal. A small amount of matrix calculation will show that $d(x, \xi) = d_n(x, \xi) + 1$ so that one may show the equivalence of the D_n and D-optimal designs using the quantities $d_n(x, \xi)$ and $d(x, \xi)$.

At the opposite extreme where $s = 1$, the variance of the least squares estimator of only the highest coefficient β_n is being minimized and Theorem 4.1 shows the support of ξ_{n-1} to be on the zeros of $(1-x^2)T_n'(x) = 0$. These zeros are $x_v = \cos v\pi/n, v = 0, 1, \cdots, n$. For the interior zeros $U_{2n-2}(x) = 1$ which implies weight $1/n$ is on each interior point. This leaves $1/2n$ on ± 1. This is the design mentioned in the introduction.

It is easily seen that $U_{2r}(\pm 1) = 2r + 1$ implying that the weight given to ± 1 by ξ_s is

$$\xi_s(+1) = \xi_s(-1) = \frac{1}{2n - s + 1} \qquad s = 1, 2, \cdots, n. \tag{5.1}$$

Whenever n is even $x = 0$ is in the support of ξ_s for each s. Since $U_{2r}(0) = (-1)^r$ we find

$$\xi_s(0) = \frac{2}{2n + 1 + (-1)^{n-s}} \qquad s = 1, 2, \cdots, n, n \text{ even}. \tag{5.2}$$

Consider the case $n = 4$ and $s = 2$, where, for example, we might have a quadratic regression but are guarding against terms involving x^3 and x^4. To investigate these terms a design using ξ_2 might be appropriate. Theorem 4.1 shows the interior support points are the zeros of

$$\rho_2'(x)\tau_3'(x) - \tfrac{1}{6}\tau_2'(x) = 0. \tag{5.3}$$

Checking the polynomials P_n and T_n, say from Davis (1963) page 369–371, equation (5.3) becomes $12x^3 - 5x = 0$. Using (5.1) and (5.2) and the symmetry of ξ_2 we find that the D_2-optimal design ξ_2 concentrates mass

$$\tfrac{1}{7}, \tfrac{9}{35}, \tfrac{1}{5}, \tfrac{9}{35}, \tfrac{1}{7} \tag{5.4}$$

on the corresponding points

$$-1, -\left(\tfrac{5}{12}\right)^{1/2}, 0, \left(\tfrac{5}{12}\right)^{1/2}, 1. \tag{5.5}$$

We can reverify that this is the D_2-optimal design by checking that $d_2(x, \xi_2) \leqslant 2$. Using the design above or using the results of Section 3, we find that the moments of ξ_2 are $c_{2i+1} = 0$, and

$$c_2 = \tfrac{1}{2}, c_4 = \tfrac{3}{8}, c_6 = \tfrac{31}{96}, c_8 = \tfrac{347}{1152}.$$

It then follows that

$$M_{21}M_{11}^{-1} = \begin{bmatrix} 0 & \frac{3}{4} & 0 \\ -\frac{1}{6} & 0 & \frac{13}{12} \end{bmatrix}$$

and

$$M_{22} - M_{21}M_{11}^{-1}M_{12} = \begin{bmatrix} \frac{1}{24} & 0 \\ 0 & \frac{1}{72} \end{bmatrix}.$$

The inequality $d_2(x, \xi_2) \leqslant 2$ then becomes

$$24\left(x^3 - \tfrac{3}{4}x\right)^2 + 72\left(x^4 - \tfrac{13}{12}x^2 + \tfrac{1}{6}\right)^2 \leqslant 2.$$

This can be checked and equality shown to hold for the support points given in (5.5).

REFERENCES

[1] DAVIS, PHILIP J. (1963). *Interpolation and Approximation*. Blaisdell, Waltham, Mass.

[2] GUEST, P. G. (1958). The spacing of observations in polynomial regression. *Ann. Math. Statist.* **29** 294–299.

[3] HOEL, P. G. (1958). Efficiency problems in polynomial regression. *Ann. Math. Statist.* **29** 1134–1145.

[4] JOLLEY, L. B. W. (1961). *Summation of Series*. Dover, New York.

[5] KARLIN, S. AND STUDDEN, W. J. (1966). *Tchebycheff Systems: With Applications in Analysis and Statistics*. Interscience, New York.

[6] KIEFER, J. (1961). Optimum designs in regression problems, II. *Ann. Math. Statist.* **32** 298–325.

[7] KIEFER, J. (1962). An extremum result. *Canad. J. Math.* **14** 597–601.

[8] KIEFER, J. and WOLFOWITZ, J. (1959). Optimum designs in regression problems. *Ann. Math. Statist.* **30** 271–294.

[9] KIEFER, J. and WOLFOWITZ, J. (1960). The equivalence of two extremum problems. *Canad. J. Math.* **12** 363–366.

[10] PERRON, O. (1954). *Die lehre von den kettenbruchen, Band I*. B. G. Teubner, Stuttgart.

[11] SEALL, ROBERT and WETZEL, MARION (1959). Some connections between continued fractions and convex sets. *Pacific J. Math.* **9** 861–873.

[12] SKIBINSKY, MORRIS (1967). The range of the $(n + 1)$th moment for distributions on [0, 1]. *J. Appl. Probability* **4** 543–552.

[13] SKIBINSKY, MORRIS (1968). Extreme nth moments for distributions on [0, 1] and the inverse of a moment space map. *J. Appl. Probability* **5** 693–701.

[14] SKIBINSKY, MORRIS (1969). Some striking properties of binomial and beta moments. *Ann. Math. Statist.* **40** 1753–1764.

[15] SZEGO, G. (1959). *Orthogonal Polynomials, rev. ed.* Amer. Math. Soc. Coll. Pub., Vol. 23, New York.

[16] WALL, H. S. (1940). Continued fractions and totally monotone sequences. *Trans. Amer. Math. Soc.* **48** 165–184.

[17] WALL, H. S. (1948). *Analytic Theory of Continued Fractions*. Van Nostrand, New York.

DEPARTMENT OF STATISTICS
MATHEMATICAL SCIENCES BUILDING
PURDUE UNIVERSITY
WEST LAFAYETTE, INDIANA 47907