_79-2_

D_s-Optimal Designs for Polynomial
Regression using Continued Fractions

by

W. J. Studden[1]
Purdue University

Department of Statistics
Division of Mathematical Science
Mimeograph Series #79-2

January 1979

D_s-Optimal Designs for Polynomial

Regression using Continued Fractions

By W. J. Studden[1]
Department of Statistics
Purdue University
W. Lafayette, IN.   47907

Abbreviated Title: D_s-Optimal Polynomial Regression

Summary. 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 n-s coeffi-

cients. The designs are calculated using continued fractions.

---

1. <u>Introduction</u>. Consider a polynomial regression situation on $[-1,1]$.
For each x or "level" in $[-1,1]$ an experiment can be performed whose out-
come 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, $i=0,1,\ldots,n$ and σ^2 are unknown. An
experimental design is a probability measure ξ on $[-1,1]$. If N observa-
tions are to be taken and ξ concentrates mass $\xi(i)$ at the points x_i,
$i=1,2,\ldots,r$ and $\xi(i)N=n_i$ are integers the experimenter takes N uncorrelated
observations n_i at each x_i, $i=1,2,\ldots,r$. 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 approxima-
tion would be needed in applications.

Let $f'(x) = (1,x,x^2,\ldots,x^n)$ 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. 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 $[-1,1]$.

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

The purpose of this paper is to consider the D_s-optimal design which minimizes the determinant of the covariance of the least squares estimates of the parameters $\beta_{s+1}, \beta_{s+2}, \ldots, \beta_n$. The estimation of all coefficients is the D-optimal situation. It will be easily seen below that this is what is referred to here as the D_o-optimal design.

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

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

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

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

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

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

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

(1.4)
$$d_s(x, \xi_s) \leq n-s.$$

To find the maximum of $|\Sigma(\xi)|$ we use the result that $|\Sigma(\xi)| = |M(\xi)||M_{11}(\xi)|^{-1}$. Note that s=0 corresponds to the D-optimal case since $M_{11}=1$. The quantity $d_0(x,\xi)$ in (1.4) is not, however, equal to $d(x,\xi)$ defined above.

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 $|\sum(\xi)|$. The relationship between the ordinary moments and the canonical moments is described in section 3. This relationship involves some simple recursive formula 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 invariance 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 polynomial are usually given on this interval and symmetry considerations allow us to calculate our examples somewhat more readily.

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

2. <u>Maximization</u> <u>of</u> $|\sum(\xi)|$.

The maximization of $|\sum(\xi)|$ 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

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

For a given set of moments $c_0, c_1, \ldots, c_{i-1}$ 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 $c_0, c_1, \ldots c_{i-1}$.
Similarly let c_i^- denote the corresponding minimum. The canonical moments are
defined by

(2.1)
$$p_i = \frac{c_i - c_i^-}{c_i^+ - c_i^-} \qquad i=1,2,\ldots$$

Note that $0 \le p_i \le 1$. 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 $p_i = 1/2$ for i odd and

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

for i even. Note that for the arc-sin law or the measure corresponding to
the ordinary Chebyshev polynomials the canonical moments are $p_i \equiv 1/2$.

<u>Theorem 2.1</u> The determinant $|\sum(\xi)|$ in (1.2) is maximized by the unique
design ξ whose canonical moments are given by

$$p_i = 1/2 \qquad \text{for i odd}$$

(2.2)
$$P_{2i} = \begin{cases} \frac{1}{2} & i=1,2,\ldots,s \\ \frac{n-i+1}{2n-2i+1} & i=s+1,\ldots,n-1 \\ 1 & i=n \end{cases}$$

<u>Proof.</u> Let Δ_{2m} denote the determinant

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

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 determinants Δ_{2m} for the moments b_i is given by

$$(2.3) \qquad \prod_{i=1}^{m} (n_{2i-1} n_{2i})^{m+1-i}$$

where

$$(2.4) \qquad n_0 = q_0, \quad n_j = q_{j-1} p_j \quad j=1,2,\ldots$$

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 multiple of 2 times the quantity (2.3). Therefore, the maximization of the determinant $|\sum(\xi)|$ is equivalent to maximizing

$$\frac{\prod_{i=1}^{n} (n_{2i-1} n_{2i})^{n+1-i}}{\prod_{i=1}^{s} (n_{2i-1} n_{2i})^{s+1-i}}$$

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, $i=0,\ldots,2n$ uniquely determine $c_0, c_1, \ldots, c_{2n}$ and the value $P_{2n} = 1$ implies that $(c_0, c_1, \ldots, c_{2n})$ 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

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

has a continued fraction expansion of the form

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

where the quantities η_i are given in (2.4). 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 $S_{oj} \equiv 1$, $j=0,1,2,\ldots$ and for $i \leq j$

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

The corresponding moments b_i are given by

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

These are taken from Skibinsky [1969] and presumably originate with works of Stieltjes. 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 $p_i = 1/2$. The powers series

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

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

$$d_i = 4m_i(1-m_{i-1})\ell_i(1-\ell_i) \qquad i=1,2,\ldots$$

(3.6) $$e_i = 1-2m_{i-1}(1-\ell_{i-1})-2(1-m_{i-1})\ell_i \qquad i=1,2,\ldots$$

$$m_i = p_{2i} \text{ and } \ell_i = p_{2i-1} \qquad i=1,2,\ldots$$

For the p_i given in (2.2) the odd $p_{2i-1} = \frac{1}{2}$ so that $e_{2i-1} = 0$ and $d_i = q_{2i-1}p_{2i} = n_{2i}$. 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 note previously that for this measure $p_{2i+1} = 1/2$ and $p_{2i} = i/(2i+1)$. Therefore we have

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

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

$$b_1 = p_1$$

$$b_2 = p_1(p_1+q_1p_2)$$

$$b_3 = p_1[p_1(p_1+q_1p_2)+q_1p_2(p_1+q_1p_2+q_2p_3)]$$

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

$$c_2 = p_2$$

$$c_4 = p_2(p_2+q_2p_4)$$

$$c_6 = p_2[p_2(p_2+q_2p_4)+q_2p_4(p_2+q_2p_4+q_4p_6)]$$

This gives $c_2 = 1/3$, $c_4 = 1/5$, $c_6 = 1/7$.

4. The Moments c_i and the Design ξ.

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)
$$p'_{n-s}(x)t'_{s+1}(x) - \alpha_s p'_{n-s-1}(x)t'_s(x) = 0$$

where

(4.2)
$$\alpha_s = \frac{1}{2} \cdot \frac{(n-s+1)}{(2n-2s+3)} \qquad s=0,1,\ldots,n-1$$

and p_i and t_i are the Legendre and Chebyshev polynomials $P_n(x)$ and $T_n(x)$ normalized so that the leading coefficient is one.

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

(4.3)
$$\xi_s(i) = \frac{2}{2n+1+U_{2s}(x_i)} \qquad i=0,1,\ldots,n$$

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

The support points of the D_s-optimal design on the polynomial in (4.1) can be expressed in another form. It is known that the support points other than ± 1 are the zeros of the polynomial orthogonal to $(1-x^2)d\xi_s(x)$. This is explained more fully below. This polynomial can be obtained by calculating the moments c_i from the p_i as explained in section 3. The required polynomial is then

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

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 $d_s(x_i,\xi_s) = n-s$. The quantity $d_s(x,\xi)$ can be re-written in the form

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

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

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

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

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

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

$$n-s = \frac{1}{\xi(i)} - (1+2\sum_{j=1}^{s} T_j^2(x_i))$$

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

$$\sum_{j=1}^{s} T_j^2(x) = \frac{1}{2}(s + \frac{\cos(s+1)\theta\sin s\theta}{\sin\theta})$$
$$= \frac{1}{2}(s - \frac{1}{2} + \frac{U_{2s}(x)}{2}).$$

<u>Proof of Theorem 4.1</u>. For the optimal design ξ_s the canonical moments $p_{2n} = 1$ implies that c_{2n} has a maximum value given the set $c_0, c_1, \ldots, c_{2n-1}$. The support points of ξ_s are then the zeros of the polynomial 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 $c_0, c_1, \ldots$ corresponding to some ξ and take

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

in a continued fraction form

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

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), p. 55.

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

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

where $d_i = m_i(1-m_{i-1})$, $i=1,2,\ldots,m_0 = 0$ 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

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

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

(4.9)
$$\sum_{i=0}^{\infty} \frac{c_{2i} - c_{2i+2}}{w^{2i+1}}$$

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

$$(4.10) \qquad \frac{c_0 - c_2}{w} \; \dot{-} \; \frac{\alpha_1}{w} \; - \; \frac{\alpha_2}{w} \; - \; \cdots \; - \; \frac{\alpha_k}{w} \; - \; \cdots$$

A number of comments are required concerning the quantities $\alpha_i = p_{2i}(1-p_{2i+1})$. For the Chebyshev or arc-sin measure proportional to $(1-x^2)^\alpha$, with $\alpha = -1/2$, the canonical moments are all $p_i \equiv 1/2$. (See Skibinsky (1969)) These are the same as the first part of the canonical moments corresponding to the optimal ξ_s. For the uniform measure $p_{2i} = i/(2i+1)$. It can readily be seen that the corresponding values $\alpha_1', \alpha_2', \ldots$ are the values $\alpha_{n-2}, \alpha_{n-3}, \ldots$ That is $\alpha_i' = \alpha_{n-1-i}$. Thus the values $\alpha_1, \alpha_2, \ldots, \alpha_{s-1}$ correspond to the Chebyshev measure, α_s is given by (4.2) and $\alpha_{s+1}, \alpha_{s+2}, \ldots, \alpha_{n-2}$ correspond to the uniform measure in reverse order. We note also that 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. For example, if we truncate the expansion (4.10) at $k \leq s-1$ the corresponding polynomial is $T_{k+2}'(w)$ with leading coefficient equal one.

The proof can now be completed by taking the terms in (4.10) up to $k = n-2$. The resulting polynomial can be seen to be the one in (4.1) using certain basic formula in continued fractions. These are given in Perron (1954) and are as follows. Let

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

and define

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

Then

$$(4.11) \quad K \begin{pmatrix} & a_2 & a_3 & \cdots & a_v \\ b_1 & b_2 & \cdots & & b_v \end{pmatrix} = K \begin{pmatrix} & a_v & a_{v-1} & \cdots & a_2 \\ b_v & b_{v-1} & & \cdots & b_1 \end{pmatrix}$$

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

$$(4.12) \quad B_{v+\lambda-1} = B_{\lambda-1}B_{v,\lambda-1} + a_\lambda B_{\lambda-2}B_{v-1,\lambda}$$

The polynomial in (4.1) is obtained from (4.12) if we let $\lambda = s+1$,

$v = -s+1$, $a_{i+1} = -\alpha_i$, $b_i = w$ and use (4.11) on the two B terms with double

subscripts.

5. Examples.

From Theorems 4.1 and 4.2 the D_0-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_0 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 $d(x,\xi) = f'(x)M^{-1}(\xi)f(x) \leq n+1$ while by (1.4) $d_0(x,\xi) \leq n$ for the D_0-

optimal. A small amount of matrix calculation will show that $d(x,\xi) = d_0(x,\xi)+1$

so that one may show the equivalence of the D_0 and D-optimal designs using the

quantities $d_0(x,\xi)$ and $d(x,\xi)$.

At the opposite extreme where $s=n-1$ and the variance of the least

squares estimator of only the highest coefficient β_n is being minimized,

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 $x_v = \cos\frac{v\pi}{n}$, $v=0,1,\ldots,n$. 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 $U_{2s}(\pm 1) = 2s+1$ implying that the weight given to ± 1 by ξ_s is

(5.1)
$$\xi_s(+1) = \xi_s(-1) = \frac{1}{n+1+s} \qquad s=0,1,\ldots,n-1$$

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

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

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

(5.3)
$$p_2'(x)t_3'(x) - \frac{1}{6} t_2'(x) = 0$$

Checking the polynomials P_n and T_n, say from Davis (1963) p.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

(5.4)
$$\frac{1}{7} \qquad \frac{9}{35} \qquad \frac{1}{5} \qquad \frac{9}{35} \qquad \frac{1}{7}$$

on the corresponding points

(5.5)
$$-1 \qquad -\sqrt{\frac{5}{12}} \qquad 0 \qquad \sqrt{\frac{5}{12}} \qquad 1.$$

We can reverify that this is the D_2-optimal design by checking that $d_2(x,\xi_2) \le 2$. 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

$$c_2 = \frac{1}{2} \qquad c_4 = \frac{3}{8} \qquad c_6 = \frac{31}{96} \qquad c_8 = \frac{347}{1152}$$

It then follows that

$$M_{22}M_{11}^{-1} = \begin{pmatrix} 0 & \frac{3}{4} & 0 \\ \\ -\frac{1}{6} & 0 & \frac{13}{12} \end{pmatrix}$$

and

$$M_{22}-M_{21}M_{11}^{-1}M_{12} = \begin{pmatrix} \frac{1}{16} & 0 \\ \\ 0 & \frac{1}{72} \end{pmatrix}$$

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

$$16(x^3 - \frac{3}{4}x)^2 + 72(x^4 - \frac{13}{12}x^2 + \frac{1}{6})^2 \leq 2.$$

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. Karlin, S. and Studden, W. J. (1966). Tchebycheff Systems: with Applications in Analysis and Statistics. Interscience, New York.

3. Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems. Canad. J. Math. 12, 363-366.

4. Kiefer, J. (1962). An extremum result, Canad. J. Math., 14, 597-601.

5. Kiefer, J. (1961). Optimum designs in regression problems, II, Ann. Math. Stat. 32, 298-325.

6. Perron, O. (1954). Die lehre von den kettenbruchen, Band I, B. G. Teubner, Stuttgort.

7. Seall, Robert and Wetzel, Marion (1959). Some connections between continued fractions and convex sets. Pacific J. Math. 9, 861-873.

8. Skibinsky, Morris (1967). The range of the (n+1)th moment for distributions on [0,1]. J. Appl. Probability 4, 543-552.

9. 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.

10. Skibinsky, Morris (1969). Some striking properties of binomial and beta moments. Ann. Math. Statist. 40, 1753-1764.

11. Szego, G. (1959). Orthogonal Polynomials, revised edition, Amer. Math. Soc. Coll. Pub., Vol. 23, New York.

12. Wall, H. S. (1940). Continued fractions and totally monotone sequences. Trans. Amer. Math. Soc. 48, 165-184.

13. Wall, H. S. (1948). Analytic theory of continued fractions. Van Nostrand, New York.