81-18

OPTIMAL DESIGNS FOR WEIGHTED POLYNOMIAL REGRESSION USING CANONICAL MOMENTS

by

W. J. Studden[1]
Purdue University, W. Lafayette, IND.

Mimeograph Series #81-18

Department of Statistics

Division of Mathematical Sciences

June 1981

[1]This research was supported by NSF Grant No. 7901707

Optimal Designs for Weighted Polynomial

Regression Using Canonical Moments

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

ABSTRACT

Consider a weighted polynomial regression of degree m 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 canonical moments.

---

[1]This research was supported by NSF Grant No. 7901707

AMS 1970 Subject classification 62K05, 62J05

Key Words. D_s-optimal, weighted polynomial regression, canonical moments.

1. Introduction. Consider a polynomial regression situation on [0,1]. For each x or "level" in [0,1] an experiment can be performed whose outcome is a random variable $y(x)$ with mean value $\sum_{i=0}^{m} \beta_i x^i$ and variance σ^2, independent of x. The parameters β_i, $i = 0,1,\ldots,m$ and σ^2 are unknown. An experimental design is a probability measure ξ on [0,1]. If N observations are to be taken and ξ concentrates mass ξ_i at the points x_i, $i=1,2,\ldots,c$ and $\xi_i N = n_i$ are integers, the experimenter takes N uncorreleated observations, n_i at each x_i, $i=1,2,\ldots,c$. 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 $m_{ij} = \int_0^1 x^{i+j} d\xi(x)$. For an arbitrary probability measure or design some approximation would be needed in applications.

Let $f'(x) = (1,x,x^2,\ldots,x^m)$ and $d(x,\xi) = f'(x)\, M^{-1}(\xi) f(x)$ when $M(\xi)$ in 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. This result holds for general regression functions. In the polynomial case the D-optimal design concentrates equal mass $(m+1)^{-1}$ on each of the m+1 zeros of $x(1-x)\, P_m'(x)$, where P_m is the mth Legendre polynomial, orthogonal to the uniform measure on [0,1]. 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 (1958)) that the design that minimizes the variance of the highest coefficient concentrates mass proportional to 1:2:2:...:2:1 on the zeros of $x(1-x)T_n'(x) = 0$ where T_n is the Chebyshev polynomial of the first kind on [0,1]. These are orthogonal with respect to $[x(1-x)]^{-1/2}$.

In the paper Studden (1980), some D_s-optimal designs where obtained. These are the designs which minimize the determinant of the covariance matrix of the least squares estimates of the highest s parameters $\beta_{r+1},\ldots,\beta_m$, where $r+s = m$. These designs where obtained using canonical moments. The D-optimal design and the design for estimating the highest coefficient are the extremal cases where $r = -1$ and $r = m-1$ respectively.

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

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

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

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

The problem of finding D_s-optimal designs is equivalent to finding the ξ maximizing the determinant of $\Sigma(\xi)$ which is given by

(1.1) $$|\Sigma(\xi)| = |M(\xi)|/|M_{11}(\xi)|.$$

We should mention the analog of the equivalence theorm for D-optimal designs mentioned in the second paragraph above. For the D_s-optimal situation the design maximizing $|\Sigma(\xi)|$ also minimizes the supremum over [0,1] of

(1.2) $$d_s(x,\xi) = (f_2 - A(\xi)f_1)'\Sigma^{-1}(f_2 - A(\xi)f_1) = f'M^{-1}f - f_1'M_{11}^{-1}f_1$$

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

(1.3) $$d_s(x,\xi_s) \leq s.$$

In the paper by Karlin and Studden (1966) ordinary D-optimal designs for certain weighted polynomial regression problems were considered. For latter reference we restate here part of the results of Theorem 5.1 of the above paper.

Theorem 1.1 Let $f'(x) = (w(x))^{1/2}(1,x,\dots,x^m)$ where $w(x) = x^{\alpha+1}(1-x)^{\beta+1}$, $x\in[0,1]$, $\alpha>-1$ and $\beta>-1$. Then the determinant $|M(\xi)|$ is uniquely maximized by the measure ξ concentrating equal mass $(m+1)^{-1}$ at the m+1 zeros of the "Jacobi" polynomial $P_{m+1}(x) = 0$. The sequence $\{P_k\}$ is orthogonal on $[0,1]$ to $x^\alpha(1-x)^\beta$.

The theorem as originally stated has other parts referring to infinite intervals. The present methods can be modified to yield these results, however, these will not be given here. In section 3 of this paper Theorem 1.1 will be obtained using canonical moments. The case $\alpha = \beta = -1$ is the D-optimal situation. The cases $\alpha = -1$, $\beta > -1$ (and $\alpha > -1$, $\beta = -1$) are considered in Theorem 3.1. In section 4, the full set of D_s-optimal designs for $w(x) = x$, $(1-x)$ and $x(1-x)$ are given analogous to the case $w(x) = 1$. The case of estimating the highest coefficient for these special $w(x)$ is given explicitly in Theorem 4.3. The full set of D_s-optimal designs for general α and β seems to involve some unresolved difficulties. In Section 2 the canonical moments are introduced and a number of technical lemmas are stated. The proofs of some of these lemmas are somewhat difficult and complete details will be given elsewhere.

2. Canonical Moments and Technical Lemmas. The original problem of finding the D-optimal design for polynomial regression is to maximize, over the

design ξ, the determinant of $M(\xi)$ where $m_{ij} = c_{i+j} = \int x^{i+j} d\xi(x)$. The usual proof involves showing that the optimal ξ is supported on m+1 points and that the weights are all equal. The determinant is then differentiated with respect to the design points which are interior to (0,1). The interior points are shown to be the roots of a polynomial which is the solution of a certain second order differential equation. The polynomial belongs to a system of orthogonal polynomials and is recognized in the original case $w(x) = 1$ (and also for the "Jacobi" case) by the differential equation. Systems of orthogonal polynomials satisfying second order differential equations arise only in the Jacobi type cases. The canonical moment approach essentially uses a parameterization of the problem in terms of the coefficients in the three term difference equation satisfied by all systems of orthogonal polynomials. This seems to be part of the reason for some extra flexibility of the solution in the D_s-optimality case with $w(x)=1$.

For an arbitrary probability measure ξ on [0,1] let $c_k = \int_0^1 x^k d\xi(x)$. For a given finite set of moments $c_0, c_1, \ldots, c_{i-1}$ let c_i^+ denote the maximum of the ith moment $\int x^i d\xi$ over the set of all measures μ having the given set of 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 moment is defined only if $c_i^- < c_i^+$.

Lemma 2.1 The canonical moments for the "Jacobi" measure $d\xi \propto x^\alpha (1-x)^\beta \, dx$ are given by

$$p_{2k-1} = \frac{\alpha + k}{\alpha+\beta+2k}$$

(2.2) $k = 1, 2, \ldots$

$$p_{2k} = \frac{k}{\alpha+\beta+2k+1}$$

These are given in Skibinsky (1969) along with some of the other lemmas presented here and other considerations.

The special cases $\alpha=\beta=0$ and $\alpha=\beta= -1/2$ are of special interest.

Corollary 2.1 The canonical moments corresponding to Lebesque measure ($\alpha=\beta=0$) are $p_{2k+1} \equiv 1/2$ and $p_{2k} = \frac{k}{2k+1}$. The case $\alpha=\beta= -1/2$ corresponding to the "arc-sin law" has $p_i \equiv 1/2$.

Many problems can be formulated very simply in terms of the canonical moments p_i and the solution found. For example we will show if m = 2 in the ordinary regression case with w(x) = 1 we have

(2.3) $$|M(\xi)| = (p_1q_1p_2)^2 (q_2p_3q_3p_4), \qquad q_i = 1 - p_i$$

The maximum of this is given simply by $p_1=p_3 = 1/2$, $p_2 = 2/3$ and $p_4 = 1$. The remainder of the solution involves converting either to the corresponding moments c_i or, more importantly, to the support points and weights in the corresponding measure ξ. This is the purpose of the majority of the lemmas. Let $1 = q_0 = \zeta_0 = \gamma_0$ and define

(2.4) $$\zeta_i = q_{i-1}p_i \text{ and } \gamma_i = p_{i-1}q_i \quad i = 1,2,\ldots$$

Lemma 2.2 If $S_{oj} = 1$, $j = 0,1,2,\ldots$ and

(2.5) $$S_{ij} = \sum_{k=i}^{j} \zeta_{k-i+1} S_{i-1k} \quad i>j$$

then $c_m = S_{mm}$.

The first few moments are

$$c_1 = p_1 = \zeta_1$$

$$c_2 = p_1(p_1+q_1p_2) = \zeta_1(\zeta_1+\zeta_2)$$

$$c_3 = \zeta_1[\zeta_1(\zeta_1+\zeta_2) + \zeta_2(\zeta_1+\zeta_2+\zeta_3)]$$

Skibinsky (1969) has shown that if the canonical moments are defined relative to any interval that they are invariant under simple linear transformations. He also shows that symmetry of the distribution is related to $p_{2i+1} = 1/2$.

Lemma 2.3 If ξ' is the measure corresponding to ξ by reversing the interval $[0,1]$, ie. letting $y = 1-x$ then $p'_{2i} = p_{2i}$ and $p'_{2i+1} = q_{2i+1} = 1-p_{2i+1}$.

The determinants that we use are one of the following forms. Let

$$\underline{\Delta}_{2m} = |c_{i+j}|_{i,j=0}^{m} \qquad \underline{\Delta}_{2m+1} = |c_{i+j+1}|_{i,j=0}^{m}$$

$$\overline{\Delta}_{2m} = |c_{i+j-1}-c_{i+j}|_{i,j=1}^{m} \qquad \overline{\Delta}_{2m+1} = |c_{i+j}-c_{i+j+1}|_{i,j=0}^{m}$$

Lemma 2.4 The above determinants are given in terms of the canonical moments by

$$\underline{\Delta}_{2m} = \prod_{i=1}^{m}(\zeta_{2i-1}\zeta_{2i})^{m+1-i} \qquad \underline{\Delta}_{2m+1} = \prod_{i=0}^{m}(\zeta_{2i}\zeta_{2i+1})^{m+1-i} \tag{2.5}$$

$$\overline{\Delta}_{2m} = \prod_{i=1}^{m}(\gamma_{2i-1}\gamma_{2i})^{m+1-i} \qquad \overline{\Delta}_{2m+1} = \prod_{i=0}^{m}(\gamma_{2i}\gamma_{2i+1})^{m+1-i}$$

The canonical moments are intimately related to orthogonal polynomials and continued fractions. As usual we use the notation

$$\frac{a_1}{b_1} + \frac{a_2}{b_2} + \frac{a_3}{b_3} = \cfrac{a_1}{b_1 + \cfrac{a_2}{b_2 + \cfrac{a_3}{b_3}}}$$

If

$$\frac{a_1}{b_1} + \frac{a_2}{b_2} + \ . \ . \ . \ \frac{a_k}{+b_k} = \frac{A_k}{B_k} \ ,$$

then an induction argument shows that A_k and B_k can be computed recursively by

$$A_k = b_k A_{k-1} + a_k A_{k-2}$$

$$B_k = b_k B_{k-1} + a_k B_{k-2}$$

One can then see that the B_k can be expressed as a tridiagonal matrix

$$B_k = \begin{vmatrix} b_1 & -1 & & & & \\ a_2 & b_2 & -1 & & 0 & \\ & a_3 & b_3 & -1 & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & 0 & & & \cdot & -1 \\ & & & & a_k & b_k \end{vmatrix}$$

This follows since these determinants satisfy the same recursive relations and the same initial conditions. Further results on continued fractions can be found in Perron (1954) or Wall (1948).

Lemma 2.5 The Stieltjes transform of measure ξ has a continued fraction expansion of the form

$$(2.6) \qquad \int_0^1 \frac{d\xi(x)}{t-x} = \frac{1}{t} - \frac{\zeta_1}{1} - \frac{\zeta_2}{t} - \frac{\zeta_3}{1} -$$

Now if the measure ξ has canonical moments that "stop" with either $p_i=0$ or 1 then the measure ξ has finite support and the transform

$$\int \frac{d\xi}{t-x} = \sum \frac{\xi_i}{t-x_i}$$

The continued fraction expansion "terminates" since ζ_k or $\zeta_{k+1} = 0$ and the transform can be written as a ratio of two polynomials. The zeros of the polynomial is the denominator are the support of ξ. This is an indication of the proof of

Lemma 2.6 If $p_k = 0$ or 1 then the support of ξ is on the zeros of

$$(2\text{-}7) \qquad D(t) = \begin{vmatrix} t & -1 & & & \\ -\zeta_1 & 1 & -1 & & \\ & -\zeta_2 & t & -1 & \\ & & -\zeta_3 & 1 & \\ & & & & \ddots \end{vmatrix}$$

The tri-diagonal determinant D(t) is truncated immediately above where the first $\zeta_i = 0$.

Suppose, for example, that $p_4 = 0$; then the resulting set of canonical moments $(p_1\ p_2\ p_3,\ 0)$ has support on two points, namely the zeros of

$$(2.8) \qquad D(t) = \begin{vmatrix} t & -1 & 0 & 0 \\ -\zeta_1 & 1 & -1 & 0 \\ 0 & -\zeta_2 & t & -1 \\ 0 & 0 & -\zeta_3 & 1 \end{vmatrix} = t^2-(\zeta_1+\zeta_2+\zeta_3)t + \zeta_1\zeta_3$$

One also has that if ξ is supported by a finite number of points then $p_k = 0$ or 1 for same k.

The next lemma essentially follows from Lemmas 2.6, 2.4 and 2.3.

Lemma 2.7 If ξ has support on $x_0, x_1, \ldots, x_m$ then

$$\prod_{i=0}^{m} x_i = \zeta_1\zeta_3\cdots\zeta_{2m+1} = p_{2m+1}\prod_{i=1}^{m} p_{2i-1}q_{2i}$$

and

$$\prod_{i=0}^{m} (1-x_i) = \prod_{i=1}^{2m+1} q_i$$

The weights on the various points can be obtained in a number of ways. Explicit formula in terms of various orthogonal polynomial are given in Karlin and Studden (1966) or Ghizzetti and Ossicini (1970). These are not given here. They can also be calculated directly by setting up the linear equations involving the weights and the ordinary moments. For example in the case just considered above; if t_1 and t_2 are the roots of (2.8) then the corresponding weights ξ_1 and ξ_2 are given by solving the equations

$$\xi_1+\xi_2 = c_0 = 1$$

$$\xi_1 t_1+\xi_2 t_2 = c_1$$

The solution for the support given in (2.7) involving D(t) gives all the support points including the endpoints t = 0 and 1 if they appear. The interior roots can be given in terms of various other orthogonal polynomials. Let four sequences of polynomials $\{P_k\}, \{Q_k\}, \{R_k\}, \{S_k\}$, $k \geq 0$, be defined by taking them orthogonal to $d\xi, t(1-t)d\xi, td\xi$ and $(1-t)d\xi$ respectively. The polynomials are normalized by taking the leading coefficients one and $P_0 = Q_0 = R_0 = S_0 = 1$.

The various moment sequences with $p_k = 0$ or 1 (for the 1st time) have corresponding measure supported by the zeros of one of the polynomials just defined. These are given by

Lemma 2.8

(a) The measure corresponding to $(p_1,\ldots,p_{2k-1},0)$ is supported on the zeros of $P_k(t) = 0$.

(b) $(p_1,\ldots,p_{2k-1},1)$ is supported by the zeros of $t(1-t)\ Q_{k-1}(t) = 0$.

(c) $(p_1,\ldots,p_{2k},0)$ is supported by the zeros of $tR_k(t) = 0$.

(d) $(p_1,\ldots,p_{2k},1)$ is supported by the zeros of $(1-t)\ S_k(t) = 0$.

Lemma 2.9 The polynomials P_k, Q_k, R_k, S_k, satisfy the recursive relations $(P_{-1} = Q_{-1} = R_{-1} = S_{-1} = 0,\ \gamma_1 = 0)$

$$\begin{aligned} P_{k+1}(t) &= (t-\zeta_{2k}-\zeta_{2k+1})\ P_k(t)-\zeta_{2k-1}\zeta_{2k}P_{k-1}(t) \\ Q_{k+1}(t) &= (t-\gamma_{2k+2}-\gamma_{2k+3})\ Q_k(t)-\gamma_{2k+1}\gamma_{2k+2}Q_{k-1}(t) \\ R_{k+1}(t) &= (t-\zeta_{2k+1}-\zeta_{2k+2})\ R_k(t)-\zeta_{2k}\zeta_{2k+1}R_{k-1}(t) \\ S_{k+1}(t) &= (t-\gamma_{2k+1}-\gamma_{2k+2})\ S_k(t)-\gamma_{2k}\gamma_{2k+1}S_{k-1}(t) \end{aligned} \tag{2.9}$$

Our final lemma is very useful and illustrates some useful symmetry in certain spaces of moments.

Lemma 2.10 (i) The support of the measures corresponding to $(p_1,\ldots,p_k,0)$ and $(p_k,\ldots,p_1,0)$ are the same.

(ii) The support of the measure corresponding to $(p_1,\ldots p_k,1)$ and $(q_k,\ldots,q_1,1)$ are the same $(q_i = 1-p_i)$.

3. D-Optimality for Classical Weights This section contains a proof of Theorem 1.1 and the extension to the cases $\alpha = -1$, $\beta > -1$ and $\alpha > -1$, $\beta = -1$. Thus we are given $c_i = \int x^i w(x) d\xi(x)$ and are required to maximize the determinant with elements $m_{ij} = c_{i+j}$, $i,j = 0,1,\ldots,m$. The determinant is of size $m+1$.

The first step of the proof is to show that the support of ξ consists of precisely m+1 points and is the same as in Karlin and Studden (1966). Given that this is the case we then may easily see that if the support of ξ is $x_0,\ldots,x_m$ then

$$|M(\xi)| = \prod_{i=0}^{m} w(x_i)\xi_i F^2(x_0,\ldots,x_m) \tag{3.1}$$

$$= \prod_{i=0}^{m} w(x_i) \prod_{i=0}^{m} \xi_i F^2(x_0,\ldots,x_m)$$

$$= \prod_{i=0}^{m} w(x_i)|M_0(\xi)|$$

Here M_0 is the matrix M when $w(x) = 1$ and $F(x_0,\ldots,x_m)$ is the determinant with rows $(1,x_i,\ldots,x_i^m)$. Note that from (3.1) we can see that the D-optimal design has equal weights since $\Pi\xi_i$ occurs as a factor in (3.1).

We now take the values for $M_0(\xi)$ and $\Pi\, w(x_i)$ from Lemmas 2.4 and 2.7. For completeness we repeat the case $w(x) = 1$. In this case

$$|M_0(\xi)| = \Delta_{2m} = \prod_{i=1}^{m} (\zeta_{2i-1}\zeta_{2i})^{m+1-i}$$

This is clearly maximized by the sequence

$$p_{2i+1} = 1/2 \tag{3.2}$$

$$p_{2i} = \frac{m-i+1}{2m-2i+1} \qquad i=1,\ldots,m-1$$

$$p_{2m} = 1$$

The proof now follows from Lemmas 2.10, 2.1 and 2.8. Thus from Lemma 2.10 (ii) we reversed the sequence (3.2) and replace p_i by q_i. By Lemma 2.1 this sequence corresponds to Lebesque measure. Now use Lemma 2.8.

Theorem 1.1 can be handled in exactly the same manner using in addition Lemma 2.7. Thus the determinant $|M(\xi)|$ is now given by

$$|M(\xi)| = M_0(\xi) \prod_{i=0}^{m} w(x_i)$$

$$= \prod_{i=0}^{m} p_{2i+1}^{\alpha+m+1-i} \; q_{2i+1}^{\beta+m+1-i} \prod_{i=1}^{m} p_{2i}^{m+1-i} \; q_{2i}^{\alpha+\beta+m+2-i}$$

This is maximized for

(3.3) $$p_{2i+1} = \frac{\alpha+m+1-i}{\alpha+\beta+2(m+1-i)} \qquad i = 0,1,\ldots,m$$

$$p_{2i} = \frac{m+1-i}{\alpha+\beta+3+2(m-i)} \qquad i = 1,2,\ldots m$$

To find the corresponding support for the D-optimal design we reverse the sequence p_i using Lemma 2.10 (i) and then refer to Lemma 2.1 and recognize the support as that for the "Jacobi" case with α and β. The support is thus the zeros obtained from Lemma 2.8 (a).

The case where $\beta = -1$ and $w(x) = x^{2+1}$, $\alpha > -1$ (or $\alpha = -1$ and $\beta > -1$) can also be readily deduced. In the case $\beta = -1$, $\alpha > -1$ the support is on $x = 1$ and m interior points. The canonical moments are the same as (3.3) with $\beta = -1$. The highest moment considered now is $p_{2m+1} = 1$ so we use Lemma 2.10 (ii). This gives the "Jacobi" canonical moments with exponent α and $\beta = -1$. The resulting support is on $x = 1$ and the m zeros of the mth polynomial orthogonal to $w(x) = x^{\alpha}$. This proves the following theorem.

<u>Theorem 3.1</u> If $w(x) = x^{\alpha+1}$ then the D-optimal design for mth degree polynomial regression has equal weight on $x=1$ and the m zeros of $P_m(x) = 0$ where $\{P_k\}$ are orthogonal with respect to $w(x) = x^{\alpha}$. If $w(x) = (1-x)^{\beta+1}$,

with $\beta > -1$, the analogous result is obtained by symmetry.

4. D_s-optimal designs for $w(x) = x$, $1-x$ or $x(1-x)$ In this section we consider the estimation of the highest s coefficients when $f'(x) = \sqrt{w(x)}\,(1,x,\ldots,x^m)$. The problem is to maximize

$$(4.1) \qquad |\textstyle\sum_s(\xi)| = |M(\xi)| \;/\; |M_{11}(\xi)|$$

where the elements of M and M_{11} have the moments $\int x^i w(x)d\xi(x)$. The matrix M_{11} is of size r+1 where r = m-s.

There appears to be inherent difficulties in evaluating the determinants $|M(\xi)|$ for general "Jacobi" weight w(x) if the support of ξ is larger than the size of the matrix $M(\xi)$. Therefore, although the ξ may be on m+1 points, when taking the ratio to evaluate $|\sum(\xi)|$ in (4.1) the denominator presents some difficulty. However the cases where w(x) = x, 1-x or x(1-x) can be obtained directly from Lemma 2.4 which does not require any restriction on the support of ξ.

The result for w(x) =x is given in Theorem 4.1, the corresponding result for w(x) = 1-x is obtained by symmetry. The result for w(x) =x(1-x) is in Theorem 4.2. The special case where s = 1 and we are estimating the highest coefficient is spelled out in Theorem 4.3.

Theorem 4.1 If w(x) = x then $|\sum_s(\xi)|$ is maximized by

$$(4.2) \qquad \begin{aligned} p_{2i} &= 1/2 \qquad i = 1,\ldots,m \\ p_{2i+1} &= \begin{cases} 1/2 & i = 1,\ldots,r \\ \dfrac{m-i+1}{2(m-i+1)+1} & i = r+1,\ldots,m-1 \end{cases} \\ p_{2m+1} &= 1 \end{aligned}$$

The support of ξ corresponding to the above p_i is given by the m+1 zeros $0 < x_1 < \ldots < x_{m+1} = 1$ of $D(t) = 0$ where $D(t)$ is given by Lemma 2.6. The corresponding weights are given by

$$\xi_i = 2\left(2m+2 + \frac{\sin 2(r+1)\theta_i}{\sin \theta_i}\right)^{-1}$$

where $2x_i - 1 = \cos \theta_i$, $0 \le \theta_i \le \pi$

Theorem 4.2. If $w(x) = x(1-x)$ then $|\Sigma_s|$ is maximized by

$$p_{2i+1} = \frac{1}{2} \qquad i = 0, 1, \ldots, m$$

(4.3)
$$p_{2i} = \begin{cases} \frac{1}{2} & i=1,\ldots,m+1 \\ \frac{m+1-i}{2(m+1-i)+1} & i=r+2,\ldots,m \end{cases}$$

$$p_{2m+2} = 0$$

The support of ξ corresponding to the above p_i is given by the m+1 zeros $0 < x_1 < \ldots < x_{m+1} < 1$ of $D(t) = 0$ where $D(t)$ is given by Lemma 2.6. The corresponding weights are given by

$$\xi_i = 2\left(2m+3 - \frac{\sin(2r+3)\theta_i}{\sin \theta_i}\right)^{-1}$$

where $2x_i - 1 = \cos \theta_i$, $0 \le \theta_i \le \pi$

Sketch of Proof of Theorems 4.1 and 4.2. The expression for $|\Sigma(\xi)|$ is evaluated from Lemma 2.4 in each case. The resulting p_i values given in (4.2) and (4.3) are then seen to maximize these expressions. The zeros are taken from Lemma 2.6. The corresponding weights are obtained by a method similar to that used in Theorem 4.2 of Studden (1980) and is omitted.

In the third paragraph of Section 1 the optimal design for estimating the highest coefficient β_m when $w(x) = 1$ was given. The design in this case has weights proportional to 1:2:2:...:2:1 on the zeros of $x(1-x)T_k'(x) = 0$ where the sequence of polynomials T_k is defined on [0,1] and are orthogonal to the arc-sin law. These zeros are x_i, $i = 0, \ldots, m$ where $2x_i - 1 = \cos \frac{i\pi}{n}$. The special case r = m-1 or s = 1 in Theorem 4.1 and 4.2 results in the following theorem.

Theorem 4.3 (i) If $w(x) = x$ then the optimal design for estimating β_m has weights proportional to 2:2:...2:1 on the m+1 points x_i, $i = 0,1,\ldots,m$ where $2x_i - 1 = \cos \theta_i$ and

$$\theta_i = \frac{2i+1}{2m+1} \pi, \quad i = 0, 1, \ldots, m$$

(ii) If $w(x) = x(1-x)$ then corresponding design has equal weight on x_i, $i =, 0, 1, \ldots, m$ where $2x_i - 1 = \cos \theta_i$, $\theta_i = \frac{2i=1}{2m+2} \pi$, $i = 0, \ldots, m$ (these are the zeros of $T_{m+1}(x) = 0$).

References

1. Ghizzetti, A. and Ossicini, A. (1970). Quadrature Formalae, Academic Press, New York.

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

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

6. Kiefer, J. and Wolfowitz, J. (1959). Optimum designs in regression problems. Ann. Math. Statist. 30, 271-294.

7. Perron, O. (1954). Die Lehre von den Kettenbruchen, B. G. Teubner, Stuttgart.

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

9. Studden, W. J. (1980). D_s-optimal designs for polynomial regression using continued fractions. Ann. Statist. 8, 1132-1141.

10. Wall, H. S. (1948). Analytic Theory of Continued Fractions. Van Nostrand, New York.