Note on Some ϕ_p-Optimal Designs for Polynomial Regression

by

W. J. Studden
Purdue University

Technical Report #87-22

Department of Statistics
Purdue University

May 1987

*

Note on Some ϕ_p-Optimal Designs for Polynomial Regression

W. J. Studden *
Purdue University Statistics Dept.
W. Lafayette, IN 47907

Abstract. In a recent paper of N. Gaffke (1987) an example was given regarding the ϕ_p-optimal designs for the highest two coefficients in a one dimensional polynomial regression. The purpose of this paper is to supply a direct proof of this result using the theory of canonical moments and orthogonal polynomials.

§ 1. Introduction. Consider a simple polynomial regression model on [-1,1]. Thus for each $x \epsilon$ [-1,1] an observation may be observed with mean value $\sum_{i=0}^{m} \theta_i x^i$ and constant variance σ^2, independent of x. The parameters $\theta' = (\theta_0, \ldots, \theta_m)$ and σ^2 are unknown. An experimental design is a probability measure ξ on [-1,1]. If N uncorrelated observations are taken and ξ has mass $\xi(i) = n_i N^{-1}$ at $x_i, i = 1, \ldots, r$, then n_i observations are taken at x_i. The covariance matrix of the least squares estimates of θ is given by $(\sigma^2/n)M^{-1}(\xi)$ where $M(\xi)$ is the information matrix of the design ξ given by

$$m_{ij} = \int_{-1}^{1} x^{i+j} d\xi(x) \tag{1.1}$$

Generally speaking the design ξ is chosen to "maximize" $M(\xi)$ or "minimize" $M^{-1}(\xi)$. Amongst criteria for this minimization are Kiefer's ϕ_p-criteria, see Kiefer (1974), Eq. 4.18 or Kiefer (1975), p. 337. The function ϕ_p are the "p-means" of $M^{-1}(\xi)$ given by

$$\begin{aligned} \phi_p(M) =& \{(m+1)^{-1} \; tr \; M^{-p}(\xi)\}^{1/p} \\ & \{(m+1)^{-1} \; tr \; (\sum_{v=0}^{m+1} \lambda_v^{-p})\}^{1/p} \end{aligned} \tag{1.2}$$

* Research supported by NSF grant #DMS-8503771

where λ_v are the eigenvalues of $M(\xi)$ and $-1 < p \leq \infty$.

If one is only interested in a subset of the parameters, say the highest s parameters, then we write

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}$$

where M_{22} is $s \times s$. The information matrix regarding these parameters is given by

$$\Lambda(\xi) = \Lambda = M_{22} - M_{21}\, M_{11}^{-1}\, M_{12} \tag{1.3}$$

The corresponding ϕ_p-criterion is to minimize

$$\phi_p(\Lambda) = \{s^{-1}\, tr\ \Lambda^{-p}\}^{1/p} \tag{1.4}$$

the values $p = 0$, 1, and ∞ are usually singled out. These values correspond to det Λ^{-1}, tr Λ^{-1} and the maximum eigenvalue of Λ^{-1}, respectively.

The present paper is concerned with the case $s = 2$ for polynomial regression. The ϕ_p-optimal designs where given rather explicitly in a recent paper of N. Gaffke (1987) which considers general regression models and is concerned with "the characterizations of design optimality and admissibility for partial parameter estimation." The special case $m = 2$ was considered by Pukelsheim (1980). The result in question is stated in the following theorem. The polynomials $U_k(x)$ and $T_k(x)$ denote the usual Tchebycheff polynomials of the 1st and 2nd kind. See, e.g. Abromowitz and Stegun (eds.) (1964).

Theorem 1 (Gaffke). The ϕ_p-optimal design for the highest two coefficients in polynomial regression of degree m on [-1,1] concentrates mass at the $m + 1$ zeros $x_0 = -1 < x_1 <$

$\ldots < x_m < x_{m+1} = 1$, of

$$(1-x^2)\,(U_{m-1}(x) + \beta U_{m-3}(x)) \tag{1.5}$$

where β is the root of

$$\left(\frac{1-\beta}{2}\right)^{p+1} - \beta = 0 \qquad 0 \le \beta < 1 \tag{1.6}$$

The corresponding weights are given by

$$\xi(x_j) = (1-\beta^2)/\{(m-1)(1-\beta^2) + (1+\beta)^2 - 4\beta T_{m-1}^2(x_j)\} \tag{1.7}$$

for $j = 1, \ldots, m-1$ and

$$\xi(-1) = \xi(+1) = \frac{1}{2}(1-\beta^2)/\{(m-1)(1-\beta^2) - (1-\beta)^2\} \tag{1.8}$$

The proof of Theorem 1 as given in Gaffke (1987) is rather elaborate and ingenious and is an application of more general results concerning partial parameter estimation. The purpose of this paper is to give a more direct proof. The proof deals directly with the moments or rather the canonical moments of the design ξ. The theory of canonical moments allows us to "identify" the ϕ_p-optimal design rather quickly. The identification or equivalence with the form in Theorem 1 is then "straightforward" but somewhat algebraically involved. For the theory of canonical moments the reader is referred to Lau (1983). See also Lau and Studden (1985), Studden (1980) (1982), and Skibinsky (1968).

§ Proof of Theorem 1. In order to prove the theorem a short description of the canonical moments and a statement of some of the results is needed. For an arbitrary design ξ the information matrix $M(\xi)$, and hence also $\Lambda(\xi)$, depends on the moments

$$c_i = \int_{-1}^{1} x^i d\xi(x) \qquad i = 1, 2, \ldots, 2m.$$

The canonical moments are defined as follows. 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 probability measures μ having moments $c_1, c_2, \ldots, c_{i-1}$. Similarly let c_i^- denote the corresponding minimum. The canonical moments are defined by

$$p_i = \frac{c_i - c_i^-}{c_i^+ - c_i^-} \quad i = 1, 2, \ldots \tag{2.1}$$

Note that $0 \le p_i \le 1$. We will have $0 < p_i < 1$ whenever $c_1, c_2, \ldots, c_i$ is in the interior of the corresponding moment space. Whenever $p_i = 0$ or 1 the subsequent p_i are left undefined. As an example consider the "Jacobi" measure with density proportional to $(1+x)^\alpha\ (1-x)^\beta$ $(\alpha > -1,\ \beta > -1)$. For this measure

$$p_{2k} = \frac{k}{\alpha + \beta + 2k + 1} \qquad k > 0$$

$$p_{2k+1} = \frac{\alpha + k + 1}{\alpha + \beta + 2k + 2} \qquad k > 0$$

The uniform measure $(\alpha = \beta = 0)$ has $p_{2k+1} = 1/2,\ k \ge 0$ and $p_{2k} = k/(2k+1)$. The "arc-sin" distribution $(\alpha = \beta = -1/2)$ has $p_k = 1/2$ for all k.

Since the underlying interval is [-1,1] and $\phi_p(\Lambda)$ is convex in M we may assume that any ϕ_p-optimal design is symmetric. In this case all the odd moments of ξ are zero and $\Lambda(\xi)$ reduces to

$$\Lambda = M_{22} - M_{21}\ M_{11}^{-1}\ M_{12}$$

$$= \begin{pmatrix} a_{m-1} & 0 \\ 0 & a_m \end{pmatrix} \tag{2.2}$$

where $a_k = \int P_k^2\ (x)\ d\xi(x),\ k = m-1, m$ and $\{P_i\}$ is the sequence of polynomials, with leading coefficient one, which are orthogonal to $d\xi(x)$. In terms of the canonical moments a_{m-1} and a_m are given (for symmetric ξ) by

$$\int_{-1}^{1} P_k^2\ (x)\ d\xi(x) = \prod_{i=1}^{k} p_{2i}\, q_{2i} \tag{2.3}$$

where $q_i = 1 - p_i$.

To obtain the ϕ_p-optimal design in terms of the p_i we may now minimize

$$a_{m-1}^{-p} + a_m^{-p}$$

with respect to p_i. This leads immediately to the following lemma. For $p = \infty$ we are simply maximizing a_m since clearly $a_{m-1} > a_m$.

Lemma 2.1. The ϕ_p-optimal design ξ_p is given by $p_{2m} = 1,\ p_i = 1/2,\ i = 1, 2, \ldots,$ $2m-1,\ i \neq 2m-2$ and $p_{2m-2} = (1+\beta)/2$ where β satisfies (1.6).

Lemma 2.1 gives, in a sense, a complete solution to the ϕ_p-optimal design problem in the present situation. It is, however, relatively "straightforward" to go from the form given in Lemma 2.1 in terms of the canonical moments to the support and weights of the design ξ_p given by Gaffke in Theorem 1. The remainder of the proof is a brief description of procedure.

In the case that $p_{2m} = 1$ it is known that the corresponding measure has support at ± 1 and m-1 points on the interior (-1,1). (See e.g. Karlin and Studden 1966, Ch. 4). The

m-1 interior points are the roots of the polynomial Q_{m-1} where $\{Q_k\}$ is the sequence of polynomial orthogonal to $(1-x^2)d\xi_p$. If ξ_p is symmetric then $p_{2i+1} = 1/2$ for all i and these polynomials (with leading coefficient equal to one) are defined recursively by

$$Q_{k+1}(x) = x\,Q_k(x) - q_{2k}\,q_{2k+2}\,Q_{i-1} \quad k \geq 1 \tag{2.4}$$

where $Q_0 \equiv 1$ and $q_i = 1 - p_i$.

We now note (as remarked earlier) that the sequence $p_i = 1/2$ for all $i \geq 1$ corresponds to the "arc-sin" measure

$$d\mu_0 = \frac{dx}{\pi\sqrt{1-x^2}}$$

The corresponding orthogonal polynomials are the Tchebycheff polynomials, $T_k(x)$, of the 1st kind. The polynomials orthogonal to $(1-x^2)d\mu_0$ correspond to the Tchebycheff polynomials of the 2nd kind devoted by $U_k(x)$. ($U_k(x)$ has leading coefficient 2^k). Since ξ_p has canonical moments $p_i = 1/2$ for $i \leq 2m-3$, it follows that, for $i \leq m-2, U_i(x) = 2^i Q_i(x)$. Inserting $k = m-2$ in (2.4) we find that

$$2^{m-1}Q_{m-1}(x) = U_{m-1}(x) + \beta U_{m-3}(x) \tag{2.5}$$

Thus the support of ξ_p is on the zeros to $(1-x^2)(U_{m-1}(x) + \beta U_{m-3}(x))$ as stated in Theorem 1.

The remaining question concerns the support $\xi_p(x_j)$ given in Theorem 1. For the interior points we use the fact that

$$\xi^{-1}(x_j) = (1-x_j^2)\sum_{k=0}^{m-2}(Q_k^*(x_j))^2 \tag{2.6}$$

where Q_k^* are orthonormal with to $(1-x^2)d\xi_p(x)$. The weight at ± 1 is given by

$$\xi(\pm 1) = 2\sum_{k=0}^{m-1}(R_k^*(1))^2. \tag{2.7}$$

where R_k^* are orthonormal with respect to $(1+x)\ d\xi_p(x)$. Formula (2.6) and (2.7) are given in Karlin and Studden (1966), Ch. 4.

To convert (2.6) to (1.7) we use the fact that $\int U_k^2(x)(1-x^2)d\mu_0 = 1/2$. Then $Q_k^*(x) = \sqrt{2}\ U_k(x)$ for $k \le m-3$. Note that the normalizing factor for Q_{m-2} uses $p_{2m-2} = 1 - q_{2m-2}$. It can be shown that

$$(Q_{m-2}^*)^2 = U_{m-2}^2(x)/q_{2m-2}.$$

Using a number of trigonometric formula (2.6) will reduce to (1.7). The details of this reduction are omitted. The considerations of (2.7) are somewhat similar and are also omitted.

References

Abramowitz, M. and Stegun, I. A. (eds) (1964). Handbook of Mathematical Functions, National Bureau of Standards, *Applied Mathematics Series* **55**.

Gaffke, Norbert. (1987). Further characterizations of design optimality and admissibility for partial parameter estimation in linear regression, *Ann. Statist.* **15**, 942–957.

Karlin, S. and Studden, W. J. (1966). Tchebycheff Systems: with applications in analysis and Statistics Interscience, NY.

Kiefer, J. (1974). General equivalence theory for optimum designs (approximate theory). *Ann. Statist.* **2**, 849–879.

Kiefer, J. (1975). Construction and optimality of generalized Yonden designs. In: A Survey of Statistical Design and Linear Models. J. N. Srivastava, ed. North Holland, Amsterdam.

Lau, T. S. (1983). Theory of canonical moments and its application in polynomial regression, Parts I and II, Purdue University, Statistics Department, *Technical Reports* #83-23 and 83-24.

Lau, T. S. and Studden, W. J. (1985). Optimal designs for trigonometric and polynomial regression using canonical moments. *Ann. Statist.* **13**, 383–394.

Pukelsheim, F. (1980). On linear regression designs which maximize information. *J. Statist. Plann. Inference.* **4**, 339–364.

Skibinsky, M. (1968). Extreme n^{th} moments for distributions on [0,1] and the inverse of a moment space map. *J. Appl. Probab.* **5**, 693–701.

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

Studden, W. J. (1982). Some robust type D-optimal designs in polynomial regression. *JASA* **77**, 916–921.