

Optimal Designs for Trigonometric and
Polynomial Regression using Canonical Moments

by

Tai-Shing Lau* and W. J. Studden*
Purdue University

Technical Report #83-26

Department of Statistics
Purdue University

July 1983

*This research was supported by NSF grant No. MCS - 8200631.

Optimal Designs for Trigonometric and
Polynomial Regression using Canonical Moments

By Tai-Shing Lau* and W. J. Studden*

Dedicated to Professor J. Kiefer

Summary. Consider a trigonometric regression of order m or a polynomial regression of degree m . Explicit D_S -optimal designs are given for some subsets of the coefficients. Bayesian type optimal designs are given for various models involving the order or the degree. The designs are calculated using canonical moments.

§1 Introduction. Consider the standard regression model where for each x or level in \mathcal{X} an experiment can be performed. The outcome is a random variable $Y(x)$ with mean $\sum_{i=1}^k \beta_i f_i(x)$ and variance σ^2 , independent of x . The parameters β_i , $i=1, \dots, k$ and σ^2 are unknown while the functions f_i , $i=1, \dots, k$ are known. An experimental design is a probability measure ξ on \mathcal{X} . If N observations are to be taken and ξ concentrates mass ξ_i at the points x_i where $N\xi_i = \mu_i$ are integers, the experimenter takes N uncorrelated observations, n_i at each x_i . The covariance matrix of the LSE of the parameters β_i is given by $(\sigma^2/N)M^{-1}(\xi)$ where $M(\xi)$ is the information matrix per observation of the design ξ with elements $m_{ij} = \int f_i f_j d\xi$. For an arbitrary probability measure

*This research was supported by NSF Grant No. MCS-8200631

or design some approximation will be needed in applications.

One of the more commonly used criterion for choosing a design ξ is the D-optimality criterion which maximizes the determinant $|M(\xi)|$. This criterion was developed largely by Kiefer (1959, 1961, 1962) and Kiefer and Wolfowitz (1959, 1960) and many others. If only a subset of the parameters is of interest the corresponding design is usually called a D_s - optimal design. This corresponds to splitting the information matrix M into blocks

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

where M_{22} is $s \times s$. The parameters β are correspondingly split into $\beta = (\beta_1, \beta_2)$ where β_2 contains the parameters of interest. The lower-right block of M^{-1} is the inverse of $\Sigma = M_{22} - M_{21} M_{11}^{-1} M_{12}$. The D_s - optimal design for estimating the parameters β_2 maximizes $|\Sigma|$. Since $|M| = |M_{11}| |\Sigma|$ this corresponds to maximizing the ratio $|M|/|M_{11}|$.

When the regression function is a polynomial on an interval, say $[-1,1]$ the matrix $M(\xi)$ becomes the classical Hankel matrix with elements $m_{ij} = b_{i+j}$, where b_v are the ordinary moments $b_v = \int_{-1}^1 x^v d\xi(x)$. In the papers Studden (1980, 1981, 1982) and Lau (1983) it was shown how explicit solutions could be found for the D or D_s - optimal criterion. This was done using orthogonal polynomials and certain canonical moments. To accomplish this the determinants $|M|$ or $|M|/|M_{11}|$ were expressed in a very simple way in terms of the canonical moments, allowing obvious maximizations. The more difficult part was recovering the design for a specified set of canonical moments. This was relatively straight forward but somewhat intricate.

The present paper is a sequel to the papers mentioned in the last paragraph. Here, attention is focused both on the polynomial regression on $[-1,1]$ and the

trigonometric case. The latter situation has regression function

$$(1) \quad g = (1, \cos \theta, \dots, \cos m\theta, \sin \theta, \dots, \sin m\theta) \quad -\pi \leq \theta \leq \pi$$

Canonical moments will again be used to analyze certain aspects of trigonometric regression on the circle $-\pi \leq \theta \leq \pi$ and to show the intimate relation between the trigonometric regression on the circle and certain polynomial models on $[-1,1]$.

In Section 2 a small portion the results for the polynomials is described and reviewed. Section 3 introduces the canonical moments for the circle and discusses some simple properties and related material. Section 4 gives a different proof of the D -optimality for certain uniform designs on the circle and gives D_S - optimal designs for the cosines or sines and indicates the relationship between these two sets of functions and the classical Chebyshev polynomials of the 1st and 2nd kind. A Bayes type analysis, originating with Lauter (1974), which provides for a sort of prior to be put on the cosines and sines is also discussed in Section 4. In Section 5 the Bayesian type analysis is carried further in discussing certain robust Bayesian type designs for the order in the trigonometric model or the degree in the polynomial model. Section 6 contains a discussion of D_S - optimal designs for determining whether odd or even terms are present while the final section relates certain half angle trigonometric series to other polynomial type functions on $[-1,1]$.

§2 Polynomial Regression. Before starting a discussion of the trigonometric regression in $(1,1)$ the canonical moments for the ordinary powers are described and some material needed in latter sections is reviewed.

Let $f(x) = (1, x, \dots, x^m)$, $x \in [-1,1]$. For an arbitrary design

or probability measure ξ on $[-1,1]$ let $b_k = \int x^k d\xi(x)$, $k = 0, 1, \dots$ denote the ordinary moments of ξ . Let b_i^+ denote the maximum value of the i th moment over the class of measures with moments $b_0, b_1, b_2, \dots, b_{i-1}$. Similarly let b_i^- denote the corresponding minimum. The canonical moments are defined by

$$(2.1) \quad p_i = \frac{b_i - b_i^-}{b_i^+ - b_i^-} \quad i = 1, 2, \dots$$

Note that $0 \leq p_i \leq 1$. By convention the canonical moments are left undefined whenever $b_i^+ - b_i^- = 0$ and the sequence is terminated. It was indicated in Studden (1980) and shown in Lau (1983) and Skibinsky (1967) that the determinant $|M(\xi)|$ could be evaluated in terms of the canonical moments. This value was given by

$$(2.2) \quad D_{2m} = |M(\xi)| = k_m \prod_{i=1}^m (\zeta_{2i-1} \zeta_{2i})^{m+1-i}$$

where $k_m = 2^{m(m+1)}$, $\zeta_0 = 1$, $\zeta_1 = p_1$, $\zeta_i = q_{i-1} p_i$, $i \geq 2$ and $p_i + q_i = 1$. A simple inspection of (2.2) shows that the D -optimal design maximizing $|M(\xi)|$ has canonical moments

$$(2.3) \quad \begin{aligned} p_{2i+1} &= 1/2 & i &= 0, 1, \dots, m-1 \\ p_{2i} &= \frac{m-i+1}{2m-2i+1} & i &= 1, 2, \dots, m-1 \\ p_{2m} &= 1 \end{aligned}$$

We have used the fact that the canonical moments p_i range "independently" over $[0,1]$. The odd moments equal to $1/2$ corresponds to measures symmetric about 0. The sequence (2.3) is a basic sequence; the even moments starting from the top are simply $1, 2/3, 3/5, 4/7$, etc. It was indicated in Studden

(1980) that these are closely related to Lebesgue measure on $[-1,1]$ which has odd canonical moments equal to one-half and even canonical moments given by $1/3, 2/5, 3/7, \dots$.

Standard procedures are available for recovering the measure or design corresponding to (2.3) or any sequence where $p_n = 0$ or 1 for some n . Some of the procedures are described in Studden (1982) and Lau (1983) and will not be given here. For the D -optimal moments in (2.3) the measure ξ , as is well known, has equal mass on the roots of $(1-x^2) P_m'(x)$, where P_m is the m th Legendre polynomial orthogonal to Lebesgue measure on $[-1,1]$.

The case where estimation of only the highest s coefficients $\beta_{r+1}, \dots, \beta_m$ ($r+s=m$) is of interest was considered in Studden (1980). The canonical moments in (2.3) then change so that $p_{2i} = 1/2$ for $i = 1, 2, \dots, r$. Certain weighted regression situations were considered in Studden (1981). Here, the regression vector $f(x) = (1, x, \dots, x^m)$ is replaced by $f(x) = \sqrt{w(x)} (1, x, \dots, x^m)$. The weighted regression is easily shown to be equivalent to letting the variance depend on x through $\sigma^2(x) = \sigma^2/w(x)$. Special cases of $w(x)$ will be used in the analysis of the trigonometric regression, so the determinants corresponding to (2.2) are listed here. The proofs are given in Lau (1983). If $f(x) = \sqrt{w(x)} (1, x, \dots, x^{m-1})$ and $w(x) = 1-x^2$ the determinant (2.2) interchanges p_i and q_i . That is

$$(2.4) \quad \bar{D}_{2m} = |M(\xi)| = k_m \prod_{i=1}^m (\gamma_{2i-1} \gamma_{2i})^{m-i+1}$$

where $\gamma_1 = q_1$, $\gamma_j = p_{j-1} q_j$, $j = 2, 3, \dots$ and k_m is given in (2.2). If $f(x) = \sqrt{w(x)} (1, x, \dots, x^m)$ and $w(x) = (1-x)$ or $(1+x)$ then the corresponding determinants are respectively given by

$$(2.5) \quad D_{2m+1} = k_m \prod_{i=0}^m (\zeta_{2i} \zeta_{2i+1})^{m-i+1}$$

and

$$(2.6) \quad \bar{D}_{2m+1} = k_m \prod_{i=0}^m (\bar{\gamma}_{2i} \bar{\gamma}_{2i+1})^{m-i+1}$$

where $\lambda_m = (m+1)^2 \lambda_1$.

Some robust type D-optimal designs were considered in Studden (1982). Here one is interested in getting close to D-optimality for regression of degree r while guarding to some extent against the coefficients $\beta_{r+1}, \dots, \beta_m$ being not zero. In Lau (1983) a rather extensive investigation of the canonical moments was undertaken. These results allowed for some simplification of the proofs of earlier results and provide many new applications.

The canonical moments for the powers can be generalized to the Fourier coefficients on the circle, where analysis is actually much simpler. This is discussed in the next section.

§3. Canonical Moments For Trigonometric Functions

Here we are dealing with the vector of regression functions g given by (1.1). For certain questions it is easier to work with the complex form using

$$(3.1) \quad h = (e^{-im\theta}, e^{-i(m-1)\theta}, \dots, 1, e^{i\theta}, \dots, e^{im\theta})$$

The functions in g are simple linear combinations of those in (3.1) so that we may write $g = Sh$ where S is a nonsingular square matrix of size $2m+1$. The information matrix $M_g = \int gg' d\theta$, for a given design on $-\pi \leq \theta \leq \pi$ can then be written as

$$\begin{aligned} M_g(\sigma) &= SM_h(\sigma)S' \\ &= ST(\sigma)JS' \\ &= ST(\sigma)\bar{S} \end{aligned}$$

The matrix J has ones down the diagonal from the upper right to lower left and zero elsewhere and the bar on S denotes complex conjugate. The matrix T is the classical Toeplitz matrix

$$(3.2) \quad T = T_{2m} = (c_{i-j})_{i,j=0}^m$$

where

$$c_k = \int_{-\pi}^{\pi} e^{-ik\theta} d\sigma(\theta) \quad k = 0, \pm 1, \pm 2, \dots$$

The determinant of T_k will be denoted by $\Delta_k = |T_k|$.

It is fairly well known that a given sequence $c_0, c_1, \dots, c_{\ell+1}$ is a trigonometric moment sequence iff $\Delta_k \geq 0, k = 0, 2, \dots, \ell+1$. In this case, if $c_0, c_1, \dots, c_{\ell}$ are given the inequality $\Delta_{\ell+1} \geq 0$ provides limits on $c_{\ell+1}$. It can be shown that the value of $c_{\ell+1}$ is contained in a specific circle depending on $c_0, c_1, \dots, c_{\ell}$. For example c_1 can be anywhere in the unit circle so that the first canonical moment is given by $a_1 = c_1$. If c_1 is fixed then $\Delta_2 \geq 0$ if and only if $|c_2 - c_1^2| \leq \Delta_1^2$. Thus c_2 lies in a circle of center c_1^2 and radius Δ_1 . The second canonical moment is therefore given by

$$a_2 = \frac{c_2 - c_1^2}{\Delta_1} = \frac{\begin{vmatrix} c_1 & c_2 \\ c_0 & c_1 \end{vmatrix}}{\Delta_1}$$

In general the $(\ell+1)$ th canonical moment is given by

$$(3.3) \quad a_{\ell+1} = (-1)^{\ell} \frac{c_{\ell+1}}{\Delta_{\ell}} \quad \ell = 0, 1, 2, \dots$$

where

$$c_{\ell+1} = |c_{i-j+1}|_{i,j=0}^{\ell}$$

The $a_{\ell+1}$ are defined only as long as $\Delta_{\ell} > 0$. A discussion of the a_{ℓ} and some of the material below can be found in many sources. The best for our purposes seems to be Geronimus (1948). He shows, among many other things, that

$$(3.4) \quad \Delta_k = \prod_{i=1}^k (1 - |a_i|^2)^{k-i+1}$$

If the measure σ is symmetric about zero there is a close connection between the quantities a_ℓ defined in (3.3) and the quantities p_ℓ defined for the polynomial case in (2.1). If σ is symmetric about zero then $\int \sin k\theta d\sigma(\theta) = 0$ and the a_ℓ are real. There is a 1-1 mapping between symmetric σ on the circle and measures ξ on $[-1,1]$ defined by projecting σ on $[-1,1]$ by the mapping $x = \cos \theta$. The function $\cos k\theta = T_k(x)$ is a polynomial in $\cos \theta$ of degree k which is the classical Chekyshev polynomial of the first kind. In this case $c_k = \int_{-1}^1 T_k(x) d\xi(x)$. Using the fact that the highest coefficient of T_k is positive we can argue from the geometrical definitions of a_ℓ and p_ℓ that we have

$$(3.5) \quad a_\ell = 2p_\ell - 1$$

That is p_ℓ is the normalized distance of b_ℓ from the lower end of its range while a_ℓ is measured from the center.

Some simple properties of p_ℓ can be derived readily from more accessible properties of a_ℓ . For example if we rotate (counter-clockwise) the measure σ through angle θ_0 to give $d\mu(\theta) = d\sigma(\theta - \theta_0)$ then the resulting moments c'_k satisfy $c'_k = e^{-ik\theta_0} c_k$. Writing down the definition of a'_k one can extract factors of $e^{i\ell\theta_0}$ from various rows and columns in the determinants involved to show that

$$(3.6) \quad a'_k = e^{-ik\theta_0} a_k \quad k = 1, 2, \dots$$

Using (3.6) with $\theta_0 = \pi$ we can immediately see the result on the p_k of reversing a measure ξ on $[-1,1]$. That is, if ξ is on $[-1,1]$, and $d\xi'(x) = d\xi(-x)$ the resulting transformation on the circle rotates θ through an angle $\theta_0 = \pi$,

in which case $a_k' = (-1)^k a_k$. The corresponding p_k then satisfy $p_{2i}' = p_{2i}$ and $p_{2i+1}' = 1 - p_{2i+1}$.

In most of the applications below, determinants related to Δ_{2m} or $|M_g|$, etc. will be maximized. The resulting answers then appear in terms of the a_i or p_i . As in the polynomial case the problem arises of how to recover the resulting design σ . In most cases the problem will involve a design σ symmetric about zero in which case the corresponding ξ can be found and projected symmetrically back onto the circle. To find the ξ , as mentioned previously, we shall appeal to results in Studden (1982) or Lau (1983). The procedure in obtaining ξ is based on the fact that the support of ξ consists of the zeros of certain orthogonal polynomials which are written recursively in terms of the canonical moments. The weights or mass on the support are then obtained by solving certain linear equations. The general trigonometric case is available. In certain limiting cases where we have an infinite number of $|a_i| < 1$ the corresponding density is of some interest. For completeness we therefore describe some of these results.

Given the values c_1, c_2, \dots the canonical moments a_1, a_2, \dots are defined by (3.3). The canonical moment a_i will satisfy $|a_i| < 1$ as long as $\Delta_{\tilde{i}} > 0$. The corresponding orthogonal system of polynomials is defined by

$$(3.7) \quad P_k(z) = \frac{1}{\Delta_{k-1}} \begin{vmatrix} c_0 & c_1 & \cdots & c_{k-1} & 1 \\ c_{-1}c_0 & \cdots & c_{k-2} & z & \\ \vdots & & & \vdots & \\ \vdots & & & \vdots & \\ c_{-k} & \cdots & & z^k & \end{vmatrix}$$

These are orthogonal with respect to σ in the sense that

$$\int P_k(z) \overline{P_\ell(z)} d\sigma = \delta_{k\ell} h_k$$

where $h_k = \Delta_k/\Delta_{k-1}$ and $z = e^{i\theta}$. Define $P_k^*(z) = z^k \overline{P(z^{-1})}$ where \overline{P} denotes that only the coefficients have been changed to complex conjugate. The polynomials P_k satisfy

$$P_0(z) \equiv 1$$

$$(3.8) \quad P_{k+1}(z) = z P_k(z) - \bar{a}_{k+1} P_k^*(z) \quad k = 0, 1, \dots$$

If the sequence a_i is such that $|a_i| < 1$, $i = 1, 2, \dots, n$ and $|a_{n+1}| = 1$, the corresponding σ is unique and supported on the zeros of $P_{n+1}(z) = 0$. These roots are all distinct and all on the unit circle $z = e^{i\theta}$. This follows since $\Delta_{n+1} = 0$ and

$$\int |P_{n+1}(z)|^2 d\sigma = \Delta_{n+1}/\Delta_n = 0$$

The weights can be found in simple cases by solving certain linear equations.

There are also general formula for the weights. If z_1, \dots, z_{n+1} are the zeros of $P_{n+1}(z) = 0$ then the corresponding weight is given by

$$(3.9) \quad \frac{\Omega_{n+1}(z_k)}{2 z_k P'_{n+1}(z_k)}$$

Here P'_{n+1} denotes derivative and the sequence Ω_k is defined as in (3.8) except a_k is replaced by $-a_k$.

The simplest case of the above is when $a_k = 0$, $k = 0, 1, \dots, n$ and $|a_{n+1}| = 1$. The support of σ is then on the zeros of

$$P_{n+1}(z) = z^{n+1} - \bar{a}_{n+1} = 0$$

These are the $n+1$ roots of unity if we take $a_{n+1} = 1$. The corresponding weights can be checked, using (3.9), to be equal.

Another interesting situation arises if we consider a sequence a_1, a_2, \dots, a_n and then take the infinite sequence by letting $a_k = 0, k > n$. The corresponding σ has density given by

$$(3.10) \quad \frac{h_n}{2\pi |P_n^*(e^{i\theta})|^2}$$

where $h_n = \Delta_n / \Delta_{n-1} = \prod_{i=1}^n (1 - |a_i|^2)$

For example if a_1 is real the sequence $a_1, 0, 0, \dots$ has corresponding density

$$(3.11) \quad \frac{1 - |a_1|^2}{2\pi (1 + a_1^2 - 2a_1 \cos\theta)} \quad -\pi \leq \theta \leq \pi$$

Note the case $a_1 = 0$ gives the uniform measure. Densities of the type (3.10)

arise as limiting cases of some of the results considered in previous papers.

For example in Kiefer and Studden (1976) the problem of extrapolating to $x_0 \in [-1, 1]$ for a polynomial regression of degree m was discussed. As m becomes large it was shown that the corresponding sequence of optimal designs converged to a measure with density

$$(3.12) \quad \frac{(x_0^2 - 1)^{\frac{1}{2}}}{\pi (1 - x^2)^{\frac{1}{2}} |x_0 - x|} \quad |x| < 1$$

This can be seen to correspond to the density in (3.11) where

$$a_i = \begin{cases} x_0 - \sqrt{x_0^2 - 1} & \text{if } x_0 > 1 \\ x_0 + \sqrt{x_0^2 + 1} & \text{if } x_0 < -1 \end{cases}$$

§4 D-optimality and designs for cosines and sines.

In this section some simple design considerations are discussed for the trigonometric regression given by (1.1). The D-optimal design is well-known and is usually described as distributing at least $2m+1$ points uniformly on $[-\pi, \pi]$. The usual proof involves an argument to the effect that a rotation invariant D-optimal design must exist. The uniform measure is therefore D-optimal. Since the D-optimal design is determined only up to the values of c_k , $k=0 \dots, 2m$ and the corresponding information matrix is unique, a design is D-optimal iff $c_k=0$, $k=0, \dots, 2m$, these being the values for the uniform design. This result is also immediate from the fact that $|M_G^-|$ is proportional to

$$(4.1) \quad \Delta_{2m} = \prod_{i=1}^{2m} (1 - |a_i|^2)^{2m-i+1}$$

This is clearly minimized by $a_i = 0$, $i = 1, 2, \dots, 2m$. In view of (3.3) this is equivalent to $c_i = 0$, $i = 1, \dots, 2m$. If the next moment a_{2m+1} is specified with $|a_{2m+1}| = 1$, the corresponding design is on $2m+1$ equally spaced points with equal weight, the exact location depending on a_{2m+1} . Many non-uniform type designs can be found by the method described in Section 3. It should be noted that the class of D-optimal designs is very large, some being of a singular nature.

It is well known that D-optimal designs for degree m are also D-optimal

for any lower order. Further the D_S -optimal design for the highest s pairs of coefficients also has $a_i = 0$, $i = 1, 2, \dots, 2m$. These two statements are immediate from (4.1).

Our next interest is in obtaining D_S -optimal designs for the sines or cosines separately. In analyzing these cases considerable use is made of the relationships between the trigonometric functions and the ordinary polynomials. Experimenters often use the fact that for certain types of analysis the straight cosines series on $[0, \pi]$ or $[-\pi, \pi]$ and the ordinary polynomials on $[-1, 1]$ are equivalent. This is due to the fact that

$$(4.2) \quad (1, \cos \theta, \dots, \cos m\theta) = (1, T_1(x), \dots, T_m(x))$$

where $\cos k\theta = T_k(x)$, $x = \cos \theta$, is the Chebyshev polynomial of the first kind. Less often used are the Chebyshev polynomials of the second kind. These correspond to the functions

$$\frac{\sin (k+1)\theta}{\sin \theta} = U_k(x) \quad x = \cos \theta$$

These are also polynomials of degree k , as indicated. We thus have

$$(4.3) \quad (\sin \theta, \dots, \sin m\theta) = \pm(1-x^2)^{\frac{1}{2}} (1, U_1(x), \dots, U_{m-1}(x))$$

It is seen that these correspond to linear combinations of the functions $f = (1, x, \dots, x^m)$ and $\pm(1-x^2)^{\frac{1}{2}} (1, x, \dots, x^{m-1})$ respectively. Since the highest coefficient of $T_k(x)$ is 2^{k-1} it follows from (2.2) that the determinant of the information matrix corresponding to the vector (4.2) is given by

$$(4.4) \quad |M_c(\theta)| = 2^{(m-1)m} |M_f(\xi)| \\ = d_m \prod_{i=1}^m (\zeta_{2i-1} \zeta_{2i})^{m+1-i}$$

where $\log d_m = m^2(m^2-1) \log 2$ and the design ξ is the projection of θ onto $[-1,1]$. Similarly one has from (2.4) that

$$(4.5) \quad |M_S(\sigma)| = d_m \prod_1^m (\gamma_{2i-1} \gamma_{2i})^{m-i+1}$$

Now if σ is symmetric about zero then terms involving a product of a sine and a cosine will vanish so that the determinant of the full information matrix $|M_g|$ splits into two parts. We have thus proven the following result.

Theorem 4.1 If σ is symmetric about zero then

$$|M_g(\sigma)| = |M_C(\sigma)| |M_S(\sigma)|$$

where $|M_C(\sigma)|$ and $|M_S(\sigma)|$ are given by (4.4) and (4.5).

Corollary 4.1 The D_S -optimal design for $(1, \cos\theta, \dots, \cos m\theta)$ in the full trigonometric model has canonical moments.

$$(4.6) \quad \begin{aligned} a_i &= 0 & i \text{ odd} \\ a_{2i} &= \frac{1}{2m-2i+1} & i = 1, 2, \dots, m. \end{aligned}$$

Corollary 4.2 The D_S -optimal design for $(\sin\theta, \dots, \sin m\theta)$ has canonical moments

$$(4.7) \quad \begin{aligned} a_i &= 0 & i \text{ odd} \\ a_{2i} &= \frac{-1}{2m-2i+1} & i = 1, 2, \dots, m \end{aligned}$$

The two corollaries follow from the theorem, by showing that the design in question must be symmetric and then maximizing either $|M_C|$ or $|M_S|$, or using (2.3) and (3.5). To force the symmetry, note that $d_\mu(\theta) = d_\sigma(-\theta)$ has the same determinant as d_σ for $|M_g|$, $|M_C|$ and $|M_S|$. By convexity the symmetrized

measure then has smaller determinant.

Example 4.1 If $m = 1$ we estimate the coefficients of $1, \cos \theta$ with a design having $a_1 = 0$ and $a_2 = 1$. This has equal mass on $\theta = 0$ and π . The coefficient of the single term $\sin \theta$ is estimated with $a_1 = 0$ and $a_2 = -1$, which has equal mass on $\pm\pi/2$. If $m = 2$ the set $1, \cos \theta, \cos 2\theta$ is estimated using $a_1 = a_3 = 0$, $a_2 = 1/3$ and $a_4 = 1$. This corresponds to masses $1/3, 1/6, 1/3, 1/6$ on the values $\theta = 0, \pi/2, \pi, 3\pi/2$. For $\sin \theta, \sin 2\theta$ we use $a_1 = a_3 = 0$, $a_2 = -1/3$ and $a_4 = -1$ which has equal mass on the 4 points corresponding to $\cos \theta = \pm 1/\sqrt{3}$.

In the design problem the sines and cosines seem to be orthogonal in some sense. That is, the design σ_c which is good for the cosines has $|M_s(\sigma_c)| = 0$ so will not estimate all the sine terms. Thus the designs in Cor. 4.1 and 4.2 are only useful if it is known before experimentation that one has either a cosine series or a sine series and not terms of both kinds. If both kinds of terms are present a D-optimal design would be more appropriate. If the cosine and sine part are of different relative importance then the designs from the following theorem might be useful.

Theorem 4.2 The design σ minimizing $|M_c|^\alpha |M_s|^\beta$ where $\alpha \geq 0, \alpha + \beta = 1$ has canonical moments

$$a_i = 0 \quad i \text{ odd}$$

$$a_{2i} = \frac{2\alpha - 1}{2m - 2i + 1} \quad i = 1, 2, \dots, m$$

Note that the a_i in theorem 4.2 are a convex combination of the corresponding values from Cor. 4.1 and 4.2. The D-optimal design, of course, has $\alpha = \beta = \frac{1}{2}$. The design in Theorem 4.2 is not unique and in this respect similar to the D-optimal design. To find a concrete example we can again let $a_{2m+1} = 1$. For $m = 1$, the first two moments are $a_1 = 0, a_2 = 2\alpha - 1$. If we let $a_3 = 1$ we find a 3-point

design on $\theta_0 = 0$ and the two values θ_1 and θ_2 where $\cos \theta = -\alpha$. The corresponding weights are $\frac{\alpha}{1+\alpha}$, $\frac{1}{2(1+\alpha)}$, $\frac{1}{2(1+\alpha)}$. A simple 4-point design can be found by setting $a_3 = 0$ and $a_4 = 1$. This has mass $\alpha/2$ on θ and π and $\beta/2$ on $\pi/2$ and $3\pi/2$.

§5 Bayesian Type Designs

In this section the Bayesian type analysis used in Theorem 4.2 is developed further. This type of criterion was introduced by L auter (1974). The idea is as follows: Suppose the experimenter has different possible models for his regression function which are indexed by k . If a prior is put on the different models, say μ_k , ($\sum \mu_k = 1$) then a possible criterion for minimization might be

$$(5.1) \quad \sum \mu_k \ln |M_k(\sigma)|$$

L auter proves a Kiefer-Wolfowitz type equivalence theorem for (5.1). Thus if f_k denotes the regression vector for the k th model and $d_k(\theta, \xi) = f_k'(\theta) M_k^{-1}(\sigma) f_k(\theta)$ then σ minimizes (5.1) if and only if σ also minimizes

$$(5.2) \quad \sup_{\theta} \sum \mu_k d_k(\theta, \sigma)$$

We have used the symbols θ and σ here. The arguments, of course, are quite general.

For the trigonometric case, the solution can be easily written down using canonical moments. Let α_k and β_k denote the prior corresponding to the terms $(1, \cos \theta, \dots, \cos k\theta)$ and $(\sin \theta, \dots, \sin k\theta)$ where $k = 1, 2, \dots, m$ and

$$\sum_{k=1}^m (\alpha_k + \beta_k) = 1.$$

Theorem 5.1 The design minimizing

$$(5.3) \quad \sum_{k=1}^m \alpha_k \ln |M_{ck}(\sigma)| + \sum_{k=1}^m \beta_k \ln |M_{sk}(\sigma)|$$

has canonical moments $a_i = 0$, i odd and

$$(5.4) \quad a_{2i} = \frac{\sum_{j=1}^{m-i+1} (\alpha_{j+1-i} - \beta_{j+1-i})}{\sum_{j=1}^{m-i+1} (2j-1) (\alpha_{j+1-i} + \beta_{j+1-i})} \quad i = 1, 2, \dots, m$$

Proof Using (4.4) and (4.5) with m replaced by k the expression in (5.3) can be written as the log of products of the canonical moments. This can be shown to be maximized by (5.4).

Example 5.1 If $m = 3$ and we set $\beta_i \equiv 0$, the result can be interpreted as assigning prior $\alpha_1, \alpha_2, \alpha_3$ to the degrees one, two and three in the polynomial model on $[-1,1]$. The canonical moments in the present case are

$$a_1 = a_3 = a_5 = 0, a_6 = 1 \text{ and}$$

$$a_4 = \frac{\alpha_2 + \alpha_3}{\alpha_2 + 3\alpha_3}, \quad a_2 = \frac{\alpha_1 + \alpha_2 + \alpha_3}{\alpha_1 + 3\alpha_2 + 5\alpha_3}$$

By using $a_i = 2p_{i-1}$ the corresponding values for p_i are $p_1 = p_3 = p_5 = 0$

$p_6 = 1$ and

$$p_4 = \frac{\alpha_2 + 2\alpha_3}{\alpha_2 + 3\alpha_3}, \quad p_2 = \frac{\alpha_1 + 2\alpha_2 + 3\alpha_3}{\alpha_1 + 3\alpha_2 + 5\alpha_3}$$

The corresponding design on $[-1,1]$ can be shown to have weight $\gamma/2$ on ± 1 and

$(1-\gamma) / 2$ on $\sqrt{p_2q_4}$ where $\gamma = p_2p_4 / (q_2 + p_2q_4)$. It may be of some interest to calculate an efficiency for the Bayes design above, obtained from Theorem 5.1; comparing it to the corresponding D-optimal designs for each degree $k = 1, 2, \dots, m$. The usual D-efficiency is defined by

$$\left(\frac{|M_k(\xi)|}{\sup_{\eta} |M_k(\eta)|} \right)^{\frac{1}{k+1}}$$

Here M_k is the information matrix for degree k . The supremum in the denominator can be calculated from (2.3) and (2.2). Simple calculations show for example that if $\alpha_1 = \alpha_2 = \alpha_3 = 1/3$ then $E_1 = .816$, $E_2 = .909$, and $E_3 = .975$. The corresponding values for $\alpha_1 = 1/2$, $\alpha_2 = \alpha_3 = 1/4$ are $E_1 = .837$, $E_2 = .906$ and $E_3 = .960$.

§6 Odd and Even Terms The analysis used in Theorem 4.1 where the determinant $|M_g|$ was factored into two parts $|M_c|$ and $|M_s|$ for a measure symmetric about zero can be carried much further. For example if σ is symmetric about 0 and also about $\pi/2$, then the terms $|M_c|$ and $|M_s|$ split further into even and odd terms. The symmetry of σ about $\pi/2$ corresponds to the projected measure ξ on $[-1,1]$ being symmetric about $x = 0$. This symmetry gives the value $p_i = 1/2$ for i odd. Let the information matrices corresponding to

$$(1, \cos 2\theta, \dots, \cos 2k\theta)$$

$$(\cos \theta, \cos 3\theta, \dots, \cos (2k+1)\theta)$$

$$(\sin 2\theta, \dots, \sin 2k\theta)$$

$$(\sin \theta, \dots, \sin (2k+1)\theta)$$

be denoted respectively by $M(c, e, k)$, $M(c, o, k)$, $M(s, e, k)$ and $M(s, o, k)$. It can then be shown that

$$(6.1) \quad |M(c, e, k)| = u_k \prod_{i=1}^k (\zeta'_{2i-1} \zeta'_{2i})^{k-i+1}$$

$$|M(c, o, k)| = v_k \prod_{i=0}^k (\zeta'_{2i} \zeta'_{2i+1})^{k-i+1}$$

$$|M(s, e, k)| = u_k \prod_{i=1}^k (\gamma'_{2i-1} \gamma'_{2i})^{k-i+1}$$

$$|M(s, o, k)| = v_k \prod_{i=0}^k (\gamma'_{2i} \gamma'_{2i+1})^{k-i+1}$$

Here $\ln u_k = k^2 \ln 2$, $\ln v_k = k(k+1) \ln 2$ and ζ'_i and γ'_i are defined similar to ζ_i and γ_i by using only the p_i for even i or doubling-up the subscripts. Thus $\zeta'_1 = p'_1 = p_2$, $\zeta'_i = q'_{i-1} p'_i = q_{2i-2} p_{2i}$, $i \geq 2$ and $\gamma'_1 = q_2$ and $\gamma'_i = p_{2i-2} q_{2i}$.

Using the expressions for the odd and even parts, D_s -type optimal designs can be obtained for $2^4 - 1 = 15$ different subsets of the parameters. In each case the design must be shown to have the required symmetry. The design is then obtained by maximizing the corresponding term in (6.1). We shall illustrate this by considering only one further case. Suppose interest is in obtaining a design which would estimate the even or odd cosine terms in the full trigonometric model. This would also give the design for the coefficients of the even (or odd) powers in the ordinary polynomial model.

The canonical moments are somewhat awkward to write down due to difference occurring when m is odd or even. The same basic sequence (2.3) however reappears again. We illustrate by example. In all cases the odd canonical moments are

equal to $1/2$. The even canonical moments are listed in the following array.

		p_2	p_4	p_6	p_8	p_{10}
$m = 2$	even	$1/2$	1			
	odd	1				
$m = 3$	even	$1/2$	1			
	odd	$2/3$	$1/2$	1		
$m = 4$	even	$1/2$	$2/3$	$1/2$	1	
	odd	$2/3$	$1/2$	1		
$m = 5$	even	$1/2$	$2/3$	$1/2$	1	
	odd	$3/5$	$1/2$	$2/3$	$1/2$	1

Suppose for example that a D_S -optimal design is required for the odd coefficients in an odd degree polynomial model. We proceed by using the canonical moments in (2.3) on the even coefficients starting at the top. In reverse order the basic sequence was $1, 1/2, 2/3, 1/2, 3/5, \dots$. These are used on the even coefficients so that every second even coefficient becomes $1/2$. This is illustrated by the bottom line in the above array.

The same procedure is used for the even coefficients in an even degree regression. If the even coefficients are required for an odd degree regression, say $2m+1$, we ignore the highest degree and revert to degree $2m$. Thus the design for the even coefficients for regressions of degree 4 and 5 are the same.

§7 Odd Degree Toeplitz Forms In the trigonometric regression model, using (1.1), of degree m , an analysis of the full information matrix involves an analysis of the Toeplitz form T_{2m} given in (3.2). It is natural to ask whether

the Toeplitz forms T_{2m+1} are relevant to any regression models. This situation arises if we use the half angle terms.

$$(7.1) \quad \cos \frac{(2k+1)\theta}{2}, \sin \frac{(2k+1)\theta}{2} \quad k = 0, 1, \dots, m$$

Using the complex form of (7.1) an analysis similar to that used at the beginning of Section 3 will result in considerations of T_{2m+1} . The functions in (7.1) are also related to certain polynomials on $[-1,1]$ in the same way that the usual trigonometric functions in (1.1) were related to the Chebychev polynomials in equations (4.2) and (4.3). We use the fact that if $x = \cos \theta$ then

$$(7.2) \quad \frac{\sin [(2k+1)\theta/2]}{\sin [\theta/2]} = W_k(x), \quad \frac{\cos [(2k+1)\theta/2]}{\cos [\theta/2]} = V_k(x)$$

are polynomials of degree k in $\cos \theta$. The sequences of polynomials in (7.2) are special Jacobi polynomials orthogonal to the weight function $(1-x)^\alpha (1+x)^\beta$. The sequence V_k corresponds to $\alpha = -1/2$ and $\beta = 1/2$ while W_k corresponds to $\alpha = 1/2$ and $\beta = -1/2$. The equations corresponding to (4.2) and (4.3) show that the cosine part of (7.1) is equal to

$$(7.3) \quad \pm [(1+x)/2]^{\frac{1}{2}} (1, V_1(x), \dots, V_{m-1}(x))$$

while the sine part of (7.1) is equal to

$$(7.4) \quad \pm [(1-x)/2]^{\frac{1}{2}} (1, W_1(x), \dots, W_{m-1}(x))$$

The information matrices for (7.3) and (7.4) are proportional to those given in (2.5) and (2.6). Nearly all of the analysis in Sections 3, 4, 5 and 6 carry over to the regression functions in (7.1). These will not be discussed further here.

References

- Geronimus, Ya L. (1948). Polynomials orthogonal on the circle and their applications. Translation Series 1, Vol. 3, Am. Math. Soc., Providence, Rhode Island.
- Kiefer, J. (1959). Optimum experimental designs. J. Roy. Statist. Soc. Ser. B 21, 272-319.
- Kiefer, J. (1961). Optimum designs in regression problems, II. Ann. Math. Statist. 32, 298-325.
- Kiefer, J. (1962). An extremum result. Canad. J. Math. 14, 597-601.
- Kiefer, J. and Studden, W. J. (1976). Optimal designs for large degree polynomial regression. Ann. Statist. 4, 1113-1123.
- Kiefer, J. and Wolfowitz, J. (1959). Optimum designs in regression problems. Ann. Math. Statist. 30, 271-294.
- Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems. Canad. J. Math. 12, 363-366.
- Lau, Tai-Shing. (1983). Theory of canonical moments and its applications in polynomial regression, Parts I and II. Purdue University Technical Reports #83-23 and 83-24.
- Läuter, E. (1974). Experimental planning in a class of models. Math. Oper. Statist. 7, 51-67.
- Skibinsky, Morris (1967). The range of the $(n+1)$ st moment for distributions on $[0,1]$. J. Appl. Probability 4, 543-552.
- Studden, W. J. (1980). D_s -optimal designs for polynomial regression using continued fractions. Ann. Statist. 8, 1132-1141.
- Studden, W. J. (1981). Optimal designs for weighted polynomial regression using canonical moments. Statistical Decision Theory and Related Topics III, Academic Press.
- Studden, W. J. (1982). Some robust type D-optimal designs in polynomial regression. JASA, 77, 916-921.