

***D*-OPTIMAL DESIGNS FOR POLYNOMIAL REGRESSION WITHOUT AN INTERCEPT**

Mong-Na Lo Huang, Fu-Chuen Chang and Weng Kee Wong*

*National Sun Yat-sen University and University of California at Los Angeles**

Abstract. *D*-optimal designs on the intervals $[a, b]$ are determined for the homoscedastic linear model with regression function $f_k^T(x) = (x, \dots, x^k)$. Motivation, properties and peculiarities of these designs are provided. In particular, the number of support points of the optimal designs for such models depends on the values of a and b , as well as an ordered eigenvalue of certain matrix. Analytical results are derived for selected values of a and b , and where they are not available, numerically optimal designs are computed. The technique here can be used to find optimal designs on more general design intervals and extend some known results (for example, Lau (1983)). Under the model considered here lower *D*- and *G*-efficiency bounds of the *D*-optimal designs for the full polynomial model are included.

Key words and phrases: *D*- and *G*-efficiency, information matrix, Jacobi polynomial, Lagrange interpolation polynomial, Sturm-Liouville equation.

1. Motivation and Background

Much of the literature in optimal experimental design assumes the regression model is a full polynomial model of degree k , i.e. $F_k^T(x) = (1, x, \dots, x^k)$. Models with improper polynomial regression functions, by which we mean $F_k(x)$ with some missing terms, have not been well studied and appear to be neglected. While making inferences for these latter models with fixed design seem to be quite straightforward, the issues of determining an optimal design for them are less so. For example, unlike the case for $F_k(x)$, closed form description of the *D*-optimal designs for improper polynomial models are not available. Although standard algorithms can generate these optimal designs readily, we find it useful to study them analytically. As we shall argue shortly, and at the end of the paper, understanding properties of these designs have potential applications to other problems. Our work here focuses primarily on the case where we have a full polynomial model but without the intercept term. Thus, our regression function has the form $f_k^T = (x, \dots, x^k)$ and relates to the full model by $F_k^T = (1, f_k^T(x))$.

There are several references in the literature concerning improper polynomial models. Many are motivated from computation problems, such as Hahn (1977),

Hedayat, Raktoc and Talwar (1977), Hawkins (1980), and Leoni (1985) while others arise in time series modeling problems, (Kraemer (1985), Lim (1992), Perron (1991), Sharma and Ali (1992)). Casella (1983) noted from the data analysis standpoint that “The problem of deciding whether an intercept model or a no-intercept model is more appropriate for a given set of data is a problem with no simple solution”. From the design perspective, it appears that only Studden (1982) and Lau (1983) have considered seeking optimal designs for $f_k(x)$ on a certain design interval, where canonical moments are used. Our technique here can be used to find optimal designs on more general design intervals and to extend some of Studden’s and Lau’s results. Models considered here occur quite naturally. For example, Lau (1983, p. 85), considered finding the relationship between the speed x and the distance y needed to stop an automobile. Suppose the model is a polynomial of degree m , i.e. $E(y) = \beta_0 + \beta_1 x + \cdots + \beta_m x^m$ with the obvious constraint that $y = 0$ if $x = 0$. This implies β_0 and possibly other lower coefficients are zero so that the model may reduce to

$$E(y) = \beta_{s+1} x^{s+1} + \cdots + \beta_m x^m$$

for some s . This is equivalent to a heteroscedastic model with regression function $F_{m-s-1}(x)$ and efficiency function $e(x) = x^{2s+2}$ (Fedorov (1972)). Thus our results here are applicable to certain types of problems involving heteroscedasticity. Further illustration of the use of models without an intercept term is given in Lau (1983, p. 85).

Our primary goal here is to find D -optimal designs for $f_k(x)$ on an arbitrary compact interval $\Omega = [a, b]$. Assume that for each x in $[a, b]$, an experiment can be performed and the outcome is a random variable $y(x)$ with mean value $\beta^T f_k(x)$ and a common variance σ^2 . Throughout it is assumed $f_k^T(x) = (x, \dots, x^k)$, the vector of parameters $\beta^T = (\beta_1, \dots, \beta_k)$ and σ^2 are unknown. Suppose n uncorrelated observations on the response $y(x)$ are to be obtained at levels x_1, \dots, x_n . An exact design specifies a probability measure ξ on $[a, b]$ which concentrates weight p_i at distinct x_i and where np_i is an integer, $i = 1, \dots, r$. An approximate design is one where the integral constraint on all the np_i is not imposed. In the latter case, the covariance matrix of the least squares estimates of the unknown parameter vector β for the model considered here is given by $(\sigma^2/n)M^{-1}(\xi)$ where

$$M(\xi) = \int_a^b f_k(x) f_k^T(x) d\xi(x)$$

denotes the information matrix of the design ξ . In this paper, we are concerned only with approximate designs.

Following the discussion in Pukelsheim (1993, p. 64), for estimation of the full parameter vector β , a design ξ is feasible if and only if $M(\xi)$ is positive

definite. Therefore for the model $f_k(x)$, the minimum number of support points of a feasible design ξ is k and does not include $x = 0$.

An approximate design ξ^* is D -optimal if ξ^* maximizes the determinant of $M(\xi)$ among all the feasible designs ξ on $[a, b]$. It is well known that the approximate D -optimal design ξ_k^* for the model $F_k(x)$ on $[-1, 1]$ is equally supported at the zeros of $(1 - x^2)P_k'(x)$, where $P_k(x)$ is the k th Legendre polynomial (Hoel (1958)). Here and throughout, the prime denotes the derivative with respect to x . The set of the support points of ξ_k^* is denoted by $S_k^* = \{x_i^*, 1 \leq i \leq k + 1\}$, where $-1 = x_1^* < \dots < x_{k+1}^* = 1$ and $x_{m+1}^* = 0$ if $k = 2m$ for some positive integer m .

A difficulty with the D -optimal design for $f_k(x)$ is that it is not invariant under linear transformations, that is the optimal design for $f_k(x)$ on a given interval does not transform linearly when the design space is translated linearly. Accordingly, due to the scale invariance property of D -optimality and the reflexivity of the D -optimal designs when the design interval is reflected for the model considered here, we discuss, without loss of generality, the problem of finding the D -optimal design for $f_k(x)$ on the interval $[a, 1]$, $-1 \leq a < 1$. The D -optimal design for the model $f_k(x)$ on $[a, 1]$ is denoted by $\xi_{a,k}^*$.

Section 2 contains preliminaries and the next three sections discuss results under three categories: (i) $-1/(k^2 + k - 1) \leq a < 1$, (ii) $a = -1$ and (iii) $-1 < a < -1/(k^2 + k - 1)$, with a convenient dichotomy for case (i) with $-1/(k^2 + k - 1) < a < y_1^*$ and $y_1^* \leq a < 1$ and $y_1^* = (x_2^* + 1)/2$. The primary tool of analysis here is the celebrated Kiefer-Wolfowitz Equivalence Theorem (Kiefer and Wolfowitz (1960)) briefly described below and solutions of certain differential equations including Sturm-Liouville equations.

2. Preliminaries

Given a design space Ω and a known regression function $f_k(x)$, the standardized variance of the estimated response at a point x under a feasible design ξ is

$$d(x, \xi) = f_k^T(x)M^{-1}(\xi)f_k(x). \quad (2.1)$$

This variance function plays a crucial role in optimal design theory as can be seen from the celebrated Kiefer-Wolfowitz Equivalence Theorem mentioned above, henceforth abbreviated as the KWT. This cornerstone result in essence states, under the framework considered here, that a feasible design ξ is D -optimal if and only if for all $x \in \Omega$, $d(x, \xi)$ is less than or equal to the number of parameters in the model. This important tool will be used to verify if a design is D -optimal, in the rest of this paper.

We first characterize the number of support points for optimal designs on

$[a, 1]$, $-1 \leq a < 1$ in Lemma 1. The proof is based on a direct application of the KWT and the fact that the variance function has a zero of multiplicity 2 at the point 0.

Lemma 1. *For the model $f_k(x)$, the D -optimal design is supported on*

- (i) k points (including 1) if $\Omega = [a, 1]$ and $0 \leq a < 1$,
- (ii) k points if k is even, and $k + 1$ points if k is odd, except possibly when $k = 1$ if $\Omega = [-1, 1]$. In either case, the optimal design is symmetric and includes ± 1 as support points.
- (iii) either k or $k + 1$ points if $\Omega = [a, 1]$, $-1 < a < 0$. In the latter case, both a and 1 are support points.

Note that if the D -optimal design is known to be supported on exactly k points it must have equal weights on those supports.

Lemma 2. *Suppose the D -optimal design $\xi_{a,k}^*$ is supported on $S_k = \{x_1, \dots, x_k\}$, and $x_k = 1$. Let $u(x) = \prod_{i=1}^k (x - x_i)$. Then*

- (i) if $a \in S_k$, there exists a real λ such that $u(x)$ satisfies the differential equation

$$x(x-a)(x-1)u''(x) + 2(x-a)(x-1)u'(x) = k(k+1)(x-\lambda)u(x), \quad (2.2)$$

- (ii) if $a \notin S_k$, then $u(x)$ satisfies the differential equation

$$x(x-1)u''(x) + 2(x-1)u'(x) = k(k+1)u(x). \quad (2.3)$$

Proof. Let $l_i(x)$, $i = 1, \dots, k$, be the fundamental Lagrange interpolation polynomials induced by the points of S_k ; then (2.1) can be written as $d(x, \xi_{a,k}^*) = k \sum_{i=1}^k [x^2 l_i^2(x)/x_i^2]$. By the KWT, we must have $d'(x_i, \xi_{a,k}^*) = 0$, $i = 2, \dots, k-1$ which implies that $x_i l_i'(x_i) = -1$. On the other hand, $l_i'(x_i) = u''(x_i)/[2u'(x_i)]$ (Pukelsheim (1993, p. 216)), and thus

$$x_i u''(x_i) + 2u'(x_i) = 0.$$

It follows that the polynomial $x(x-a)(x-1)u''(x) + 2(x-a)(x-1)u'(x)$ is one degree higher than that of $u(x)$ and the zeros of $u(x)$ are zeros of this polynomial as well. There is, thus, a real λ and a constant c such that

$$x(x-a)(x-1)u''(x) + 2(x-a)(x-1)u'(x) = c(x-\lambda)u(x).$$

Comparing the leading coefficients on both sides shows $c = k(k+1)$ and the result follows. For case (ii), the argument is similar.

3. Optimal Designs on $[a, 1]$, $-1/(k^2 + k - 1) \leq a < 1$

We are now ready to find the D -optimal designs for $f_k(x)$ on various subintervals $[a, 1]$, $-1 \leq a < 1$. To this end, let $y_i^* = (x_{i+1}^* + 1)/2$, $i = 1, \dots, k$.

Theorem 1. For the model $f_k(x)$ on $[a, 1]$, where $-1/(k^2 + k - 1) \leq a \leq y_1^*$, the design $\xi_{0,k}^*$ which concentrates equal weights at y_i^* , $1 \leq i \leq k$, is D -optimal.

Proof. First consider the case when $a = 0$. Since the point $a = 0$ cannot be a support point, the $u(x)$ defined in Lemma 2 satisfies (2.3). Letting $v(x) = xu(x)$, it is easy to verify that $x(x-1)u''(x) = (x-1)v''(x) - 2(x-1)u'(x)$ and the following differential equation holds

$$x(x-1)v''(x) = k(k+1)v(x). \quad (3.1)$$

It is well known that the unique polynomial solution for (3.1) up to a constant factor is the polynomial $x(x-1)P_{k-1}^{(1,1)}(2x-1)$ (Szegö (1975, 4.21.1)), where for any $n \geq 1$, $\alpha, \beta > -1$, $P_n^{(\alpha,\beta)}(x)$ is the Jacobi polynomial defined by

$$P_n^{(\alpha,\beta)}(x) = \frac{(-1)^n}{2^n n!} (1-x)^{-\alpha} (1+x)^{-\beta} \frac{d^n}{dx^n} [(1-x)^{n+\alpha} (1+x)^{n+\beta}]. \quad (3.2)$$

Since the zeros of $x(x-1)P_{k-1}^{(1,1)}(2x-1)$ are the support points of the approximate D -optimal design for $F_k(x)$ on $[0, 1]$, we have proved the case for $a = 0$.

Next consider when $a \in (0, y_1^*]$. On this interval, it is straightforward to show the variance function $d(x, \xi_{0,k}^*)$ satisfies the KWT; therefore $\xi_{0,k}^*$ is still D -optimal. Similarly for other intervals with $a < 0$, design $\xi_{0,k}^*$ remains optimal if the variance function $d(x, \xi_{0,k}^*)$ stays bounded above by k for all x in $[a, 1]$. To determine the smallest value of a , $-1 < a < 0$, for which this is true solve for the smallest zero t of $d(x, \xi_{0,k}^*) - k$. Evidently, this polynomial has double zeros at y_i^* , $1 \leq i \leq k-1$ and a single zero at $y_k^* = 1$, where y_i^* are the zeros of the derivative of the Legendre polynomial of degree k on $[0, 1]$. Accordingly, write

$$d(x, \xi_{0,k}^*) - k = cQ_{k-1}^2(x) [x^2 - (1+t)x + t],$$

where $Q_{k-1}(x) = \prod_{i=1}^{k-1} (x - y_i^*)$. But y_i^* , $1 \leq i \leq k-1$ are also the zeros of the orthogonal polynomial with respect to the weight function $x(1-x)$ on $[0, 1]$, i.e. the zeros of the Jacobi polynomial $P_{k-1}^{(1,1)}(2x-1)$; consequently,

$$Q_{k-1}(x) = \prod_{i=1}^{k-1} (x - y_i^*) = \sum_{i=0}^{k-1} q_i x^i = \sum_{i=0}^{k-1} \binom{k}{i} \binom{k}{k-1-i} (x-1)^{k-1-i} x^i.$$

The last equality can be obtained from the representation of Jacobi polynomial on $[0, 1]$ (Szegö (1975, p. 68)). In addition, $d(x, \xi_{0,k}^*) - k$ has a local minimum at $x = 0$, so the coefficient of x is 0, yielding

$$2tq_1 = (1+t)q_0.$$

Additional computation shows

$$q_0 = (-1)^{k-1}k \text{ and } q_1 = (-1)^{k-2}k(k^2 + k - 2)/2,$$

which yields

$$t = q_0/(2q_1 - q_0) = -1/(k^2 + k - 1).$$

This completes the proof of the theorem.

Now consider the case when $a \in (y_1^*, 1)$. By Lemma 1, there are k support points for the D -optimal design including the right end point 1. The left end point a is also a support point of the optimal design; otherwise from Lemma 2, $u(x) = \prod_{i=1}^k (x - x_i)$ satisfies (2.3), which would imply $\xi_{0,k}^*$ is still D -optimal, contradicting $y_1^* < a < 1$. By Lemma 2, $u(x) = \prod_{i=1}^k (x - x_i) = \sum_{i=0}^k s_i x^i$ satisfies (2.2) and comparing coefficients on both sides of (2.2) yields

$$(1 - \tau_{i-1})s_{i-1} + (a + 1)\tau_i s_i - a\tau_{i+1} s_{i+1} = \lambda s_i, \tag{3.3}$$

where $\tau_i = i(i + 1)/[k(k + 1)]$, $i = 0, \dots, k$, $\tau_{-1} = 1$ and $\tau_{k+1} = 0$.

We are now ready to state the D -optimal design for the model $f_k(x)$ when $\Omega = [a, 1]$ and $y_1^* < a < 1$.

Theorem 2. *For the model $f_k(x)$ on $[a, 1]$, $y_1^* < a < 1$, the D -optimal design $\xi_{a,k}^*$ is equally supported on the zeros of the monic polynomial $u(x) = \sum_{i=0}^k s_i x^i$, where the coefficient vector $s^T = (s_0, \dots, s_k)$ of $u(x)$ is the unique eigenvector with $s_k = 1$ corresponding to the smallest eigenvalue ($\neq a$ or 1) of the $(k + 1) \times (k + 1)$ tridiagonal matrix*

$$A = A(a) = \begin{pmatrix} (a + 1)\tau_0 & -a\tau_1 & \cdots & 0 & 0 \\ 1 - \tau_0 & (a + 1)\tau_1 & \ddots & \vdots & \vdots \\ 0 & 1 - \tau_1 & \ddots & -a\tau_{k-1} & 0 \\ \vdots & \vdots & \ddots & (a + 1)\tau_{k-1} & -a\tau_k \\ 0 & 0 & \cdots & 1 - \tau_{k-1} & (a + 1)\tau_k \end{pmatrix}, \tag{3.4}$$

where $\tau_i = i(i + 1)/[k(k + 1)]$, $i = 0, \dots, k$.

Before proving Theorem 2 we first prove a lemma.

Lemma 3. *All the eigenvalues of A defined in (3.4) are real.*

Proof. After rearranging the equations in (3.3), the coefficient vector $s^T = (s_0, \dots, s_k)$ of $u(x)$ can be expressed as an eigenvector of the matrix A as defined in (3.4) with corresponding eigenvalue λ . Note that both a and 1 are eigenvalues of A . This can be seen by substituting λ with a and 1 into (2.2) respectively.

Also note that when $\lambda = a$, (2.2) reduces to the differential equation satisfied by the unique polynomial solution of (2.3) on $[0, 1]$ discussed in Theorem 1. The case for $\lambda = 1$ can be proved similarly. Therefore the optimal design can not be obtained through eigenvalues a and 1.

Now let $\lambda_j, j = 1, \dots, r$, denote the distinct eigenvalues of A excluding a and 1. First, from the fact that A is tridiagonal with positive subdiagonal entries $1 - \tau_0, \dots, 1 - \tau_{k-1}$, it is easy to see that $s_k \neq 0$. Note that the eigenspace for each λ_j is of dimension 1, which can be seen through row operations in keeping the subdiagonal entries and eliminating all the other entries above the subdiagonal. Therefore there is unique eigenvector with $s_k = 1$ for each λ_j . Let $p(x) = x^2, q(x) = k(k+1)x^2 / [(x-a)(x-1)]$ and $\rho(x) = k(k+1)x / [(x-a)(1-x)]$. Moreover, as $\mathbf{1}^T A = \mathbf{1}^T$ and $f_k^T(a)A = a f_k^T(a)$, where $\mathbf{1}^T = (1, \dots, 1)$ and $f_k^T(a) = (a, \dots, a^k)$, it follows that for the corresponding eigenvalue λ , $u(1) = \mathbf{1}^T s = \mathbf{1}^T A s = \lambda \mathbf{1}^T s = \lambda u(1)$. Similarly $au(a) = \lambda u(a)$. Now by the fact that $\lambda \neq a$ or 1, we have $u(a) = u(1) = 0$. Then for a given λ , write (2.2) as a Sturm-Liouville equation

$$L[u] - \lambda \rho(x)u = 0, \quad u(a) = u(1) = 0, \quad (3.5)$$

where $L = D[p(x)D] - q(x)$ is a linear differential operator. Although here $q(x)$ and $\rho(x)$ are singular on the boundary points, several properties of the regular Sturm-Liouville problems still apply to the problem here. For example, the eigenvalues of (3.4) are all real. To see this, observe from (3.3) that corresponding to each λ_j in (3.4), there is a unique monic polynomial solution $u_j(x)$. Furthermore, L is a self adjoint operator for $\{u_j\}_1^r$, where the inner product is defined by

$$(u_i, u_j) = \int_a^1 u_i \bar{u}_j dx,$$

and so all the λ_j are real. For more details about the Sturm-Liouville problem, see Boyce and Diprima (1992), or Birkhoff and Rota (1989).

Proof of Theorem 2. For convenience, let $\lambda_1 < \dots < \lambda_r$. We will argue that the unique polynomial solution u_1 corresponding to the smallest eigenvalue λ_1 is the only polynomial which has all k zeros in $[a, 1]$. Then it is the unique solution from which the D -optimal design on $[a, 1], a \geq y_1^*$, is obtained.

Let n_j be the number of zeros of u_j which lie in $[a, 1]$. For any pair $\lambda_i < \lambda_j$, we first prove that $n_i > n_j$. To this end, let $\kappa_1 < \kappa_2$ be two successive zeros of $u_j(x)$ on $[a, 1]$. Without loss of generality, suppose both $u_i(x)$ and $u_j(x)$ are positive on (κ_1, κ_2) and let

$$w(x) = p(x) [u_i(x)u_j'(x) - u_i'(x)u_j(x)].$$

Then

$$w(\kappa_1) = p(\kappa_1)u_i(\kappa_1)u'_j(\kappa_1) \geq 0,$$

and

$$w(\kappa_2) = p(\kappa_2)u_i(\kappa_2)u'_j(\kappa_2) \leq 0,$$

which implies $w'(x) \leq 0$ for some $x \in (\kappa_1, \kappa_2)$. On the other hand, for $x \in (\kappa_1, \kappa_2)$

$$\begin{aligned} w'(x) &= \left(u_i(x) \left[p(x)u'_j(x) \right] - \left[p(x)u'_i(x) \right] u_j(x) \right)' \\ &= u_i(x) \left[p(x)u'_j(x) \right]' - \left[p(x)u'_i(x) \right]' u_j(x) \\ &= (\lambda_j - \lambda_i)\rho(x)u_i(x)u_j(x) > 0. \end{aligned}$$

Hence, $w(x)$ is increasing on (κ_1, κ_2) , which is a contradiction. Since both a and 1 are zeros for both u_i and u_j , we must have $n_i > n_j$. Finally, since there is at least one solution of u with k zeros in $[a, 1]$, the solution must correspond to the smallest eigenvalue λ_1 .

To illustrate the use of Theorem 2, consider the problem of finding the D -optimal designs when $k = 4$ and $\Omega = [1/2, 1]$. First note that $1/2 = a > y_1^* = (x_2^* + 1)/2 = 0.17267$, and the matrix A defined in (3.4) is

$$A = \begin{pmatrix} 0 & -1/20 & 0 & 0 & 0 \\ 1 & 3/20 & -3/20 & 0 & 0 \\ 0 & 9/10 & 9/20 & -3/10 & 0 \\ 0 & 0 & 7/10 & 9/10 & -1/2 \\ 0 & 0 & 0 & 2/5 & 3/2 \end{pmatrix}.$$

The eigenvalues of A are $(10 - \sqrt{19})/20 = 0.282055, 1/2, 1/2, (10 + \sqrt{19})/20 = 0.717945$ and 1. The smallest eigenvalue ($\neq 1/2$ and 1) of A is $(10 - \sqrt{19})/20$ and the corresponding eigenvector is $(0.292466, -1.64983, 3.40223, -3.04486, 1)$. Therefore the polynomial $u(x) = x^4 - 3.04486x^3 + 3.40223x^2 - 1.64983x + 0.292466$, and the D -optimal design is equally supported at the zeros of $u(x)$, i.e. 0.5, 0.664177, 0.880685 and 1.

4. Optimal Designs on $[-1, 1]$

Following a similar argument using the differential equation (2.2), we show that, excluding 0, the support points of the D -optimal design $\xi_{-1,k}^*$ coincide with those of the D -optimal design ξ_k^* for $F_k(x)$ on $[-1, 1]$ when k is even.

Theorem 3. *For the model $f_k(x)$ with $k = 2m$ on $[-1, 1]$, the design $\xi_{-1,k}^*$ which concentrates equal weights on the points of $S_k^* \setminus \{0\}$ is D -optimal.*

Proof. From Lemma 1 and (2.2) the polynomial $u(x)$ satisfies

$$x(x^2 - 1)u'' + 2(x^2 - 1)u' = k(k + 1)(x - \lambda)u.$$

Note that $u(x)$ is symmetric with $u'(0) = 0$ and $u(0) \neq 0$. After substituting $x = 0$ in the above equation, it follows that $\lambda = 0$. Hence $u(x)$ satisfies

$$x(x^2 - 1)u'' + 2(x^2 - 1)u' = k(k + 1)xu.$$

Now let $v(x) = xu(x)$. By a similar argument in the proof of Theorem 1 we have

$$(x^2 - 1)v'' = k(k + 1)v,$$

whose unique polynomial solution is $(1 - x^2)P_{k-1}^{(1,1)}(x)$, where the Jacobi polynomial $P_{k-1}^{(1,1)}(x)$ is as defined in (3.2) up to a constant factor, (Abramowitz and Stegun (1964, 22.6.2)). Since the zeros of $(1 - x^2)P_{k-1}^{(1,1)}(x)$ are the support points of ξ_k^* , the proof for the case $a = -1$ and even k is complete.

For the case when k is odd, the problem is apparently much harder and analytical results are not available. We computed the optimal design using a Fortran program which calls a maximization routine from IMSL. In each case, the starting design is ξ_k^* and the optimality of the design was verified using the KWT with $\max_x d(x, \xi) - k \leq 10^{-7}$. Table 1 shows the positive support points and corresponding weights of the numerically D -optimal designs for odd values of k between 3 and 15, and the negative supports can be obtained symmetrically.

Table 1. D -optimal designs for odd degree on $[-1, 1]$

3	0.602	1.000						
	0.178	0.322						
5	0.434	0.781	1.000					
	0.124	0.178	0.198					
7	0.338	0.622	0.875	1.000				
	0.097	0.123	0.138	0.142				
9	0.277	0.515	0.747	0.927	1.000			
	0.080	0.095	0.105	0.109	0.111			
11	0.234	0.439	0.645	0.823	0.945	1.000		
	0.068	0.077	0.085	0.089	0.090	0.091		
13	0.203	0.382	0.566	0.734	0.869	0.960	1.000	
	0.059	0.065	0.072	0.075	0.076	0.076	0.077	
15	0.179	0.339	0.503	0.660	0.795	0.900	0.970	1.000
	0.053	0.057	0.062	0.064	0.065	0.066	0.066	0.067

Left column: degree k of the polynomial regression without intercept.

Right columns: positive support points in upper row, weights in lower row.

It can be seen that the support points of these optimal designs are more dense near the boundaries than in the center 0. The weight also increases as $|x_i|$ increases and the optimal weight approaches $1/k$ as k increases. This is justified in Lemma 4.

Lemma 4. *Let $w_1 \leq \dots \leq w_{k+1}$ be the ordered weights of the optimal design $\xi_{-1,k}^*$. Then for $i = 1, \dots, k + 1$, the weights w_i satisfy*

$$\frac{1}{k} \left(\frac{i - 1}{i} \right) \leq w_i \leq \frac{1}{k}.$$

Proof. The upper bound for w_i is given by Pukelsheim (1993, p. 201). Note that $\sum_{j=1}^i w_j = 1 - \sum_{j=i+1}^{k+1} w_j \leq iw_i$ and $\sum_{j=i+1}^{k+1} w_j \leq (k + 1 - i)/k$. Combining these two inequalities, we obtain the desired lower bound for w_i .

5. Optimal Designs on $[a, 1]$, $-1 < a < -1/(k^2 + k - 1)$

General analytical formulas are not available for this case. However, it is known that the support points of the optimal design are either k or $k + 1$ depending on the values of a and k . Numerically optimal designs can be constructed with the aid of Table 2 below. To gain insight on how the results in Table 2 are obtained, it is instructive to consider the case $k = 2$ in some detail. Results for $k \geq 3$ are derived similarly.

Table 2. Intervals of a for D -optimal design with k support points

k
2 $(-1.000, -0.217)(-0.200 = -2/5, 1)$
3 $(-0.930, -0.104)(-0.091 = -1/11, 1)$
4 $(-1.000, -0.531)(-0.443, -0.062)(-0.053 = -1/19, 1)$
5 $(-0.903, -0.323)(-0.267, -0.041)(-0.034 = -1/29, 1)$
6 $(-1.000, -0.678)(-0.552, -0.220)(-0.180, -0.029)(-0.024 = -1/41, 1)$
7 $(-0.903, -0.465)(-0.379, -0.160)(-0.131, -0.022)(-0.018 = -1/55, 1)$
8 $(-1.000, -0.761)(-0.623, -0.342)(-0.278, -0.122)(-0.100, -0.017)(-0.014 = -1/71, 1)$
9 $(-0.908, -0.561)(-0.461, -0.263)(-0.214, -0.096)(-0.079, -0.013)(-0.011 = -1/89, 1)$
10 $(-1.000, -0.812)(-0.674, -0.433)(-0.357, -0.209)(-0.171, -0.078)(-0.064, -0.011)(-0.009 = -1/109, 1)$

Theorem 4. *For the model $f_2(x)$ defined on $[a, 1]$, there exists a number a_0 such that the design $\xi_{a,2}$ supported with equal mass on the two points a and 1, is D -optimal if $-1 < a \leq a_0$, and a_0 is the unique root in $[-1, -1/5]$ of the equation*

$$z^2[(1 - z)^2 + a^2(z - a)^2] - a^2(1 - a)^2 = 0, \tag{5.1}$$

where

$$z = z(a) = \left[3(1 + a^3) - (a^6 - 8a^4 + 18a^3 - 8a^2 + 1)^{1/2} \right] / [4(1 + a^2)].$$

Moreover, if $a_0 < a < -1/5$, the D -optimal design $\xi_{a,2}^*$ has three support points $a, -2a/(1+a), 1$ with corresponding weights $w_1, w_2, 1 - w_1 - w_2$, where

$$w_1 = 4(1 + 5a) / [(1 - a^2)(3 + a)(1 + 6a + a^2)],$$

and

$$w_2 = (-1 - 4a + 2a^2 - 4a^3 - a^4) / [(3 + a)(1 + 3a)(1 + 6a + a^2)].$$

Proof. The key idea is the KWT applied to the behavior of the variance function for $d(x, \xi_{a,2})$. Straightforward algebra verifies

$$d(x, \xi_{a,2}) = 2x^2[(1 - x)^2 + a^2(x - a)^2] / [a^2(1 - a)^2],$$

and

$$d'(x, \xi_{a,2}) = 4x[(2a^2 + 2)x^2 - 3(a^3 + 1)x + (a^4 + 1)] / [a^2(1 - a)^2].$$

The discriminant of the quadratic polynomial in the numerator of $d'(x, \xi_{a,2})$, is

$$9(1 + a^3)^2 - 4(2a^2 + 2)(1 + a^4) = a^6 - 8a^4 + 18a^3 - 8a^2 + 1.$$

One can verify that there exists a unique $a_1 \in (-1, -1/5)$, such that this discriminant vanishes at a_1 and is negative for $a \in (-1, a_1)$ and nonnegative for $a \in [a_1, 0)$ (Numerical calculation shows a_1 is approximately -0.27228). Then for $-1 < a < a_1$ $d'(x, \xi_{a,2}) < 0$ if $x < 0$, $= 0$ if $x = 0$ and > 0 if $x > 0$, which implies $d(x, \xi_{a,2})$ has no local maximum in interval $(a, 1)$ for $-1 < a < a_1$. Consequently, if $-1 < a < a_1$, $\xi_{a,2}$ is optimal by the KWT. Next, for the case $a_1 \leq a < -1/5$, algebra shows that $d(x, \xi_{a,2})$ has a local maximum in $(a, 1)$ at $z = z(a)$ defined near (5.1), and $d(z, \xi_{a,2})$ is a continuous and increasing function of a on the interval $a_1 \leq a < -1/5$. Solving the equation $d(z, \xi_{a,2}) = 2$ in terms of a yields a unique solution a_0 which satisfies (5.1) ($a_0 \approx -0.216845$). It follows that for all x , $d(x, \xi_{a,2}) \leq 2$ if $a_1 \leq a \leq a_0$, but for $a_0 < a < -1/5$ it follows that $d(z, \xi_{a,2}) > 2$. This implies a two-point design with equal weights can not be optimal for $a_0 < a < -1/5$. Therefore if $a_0 < a < -1/5$, the optimal design $\xi_{a,2}^*$ has three support points. Now, for a design ξ with three supports at $a, s, 1$ and corresponding weights $w_1, w_2, 1 - w_1 - w_2$, the optimal solutions s, w_1 and w_2 maximizing the determinant of $M(\xi)$, under the constraints that $a < s < 1, w_1 > 0, w_2 > 0$ and $1 - w_1 - w_2 > 0$, can be found by solving the

equations obtained through taking partial derivatives of the determinant of $M(\xi)$ with respect to s, w_1 and w_2 , where w_1, w_2 can be expressed easily in terms of s and later the optimal s can be determined. This yields the solutions as stated in the theorem. Then the variance function of the design $\xi_{a,2}^*$ can be shown to be

$$d(x, \xi_{a,2}^*) = 2 + (x - a)(1 - x)[2a + (1 + a)x]^2 / (2a^3) \leq 2, \quad \text{for all } x \in [a, 1],$$

where the KWT is satisfied, and the proof is complete.

For $3 \leq k \leq 10$, we also indicate in Table 2 the intervals of values of a such that the number of supports of the D -optimal design on $[a, 1]$ are k . The optimal designs $\xi_{a,k}^*$ are supported on $k+1$ points for those values a in $(-1, -1/(k^2+k-1))$, but not in any of the intervals given in Table 2. In addition, observe that the two end points are in the supports of the optimal design if $a < -1/(k^2+k-1)$. Now, using (3.3) we can determine the optimal design by exhausting all possible polynomial solutions $u(x)$ defined in (2.2) if the number of supports of the optimal design is k . It is found that there are $[k/2]+1$ intervals of a , such that the optimal design is supported by exactly k points. Here $[x]$ denotes the greatest integer smaller than or equal to x . After an extensive numerical study we have also found an interesting relation between the intervals and the eigenvalues of A defined in (3.4) corresponding to the optimal designs. Specifically, the support points of the optimal design are the zeros of the polynomial solution corresponding to the i th smallest eigenvalue if a is in the i th interval from the right in Table 2. For instance, if $k = 4$ and $a = -1/3$, a is in the second interval from the right, then the eigenvalues of A are $-1/3, -0.0643053, 0.192547, 0.538425$, and 1 , and the support points of the optimal design are $-1/3, 0.376862, 0.783901$, and 1 , which are the zeros of the polynomial $u(x) = x^4 - 1.82743x^3 + 0.735932x^2 + 0.189973x - 0.0984742$. Note that the coefficients of $u(x)$ are obtained from the eigenvector corresponding to the second smallest eigenvalue -0.0643053 . Similarly it can be verified that if $a = -2/3$, a is in the third interval from the right and the eigenvalues of A are $-2/3, -0.246218, 0.0949928, 0.484559$, and 1 . The support points of the optimal design are $-2/3, -0.417435, 0.679953, 1$ which are the zeros of $u(x) = x^4 - 0.595851x^3 - 0.862997x^2 + 0.269624x + 0.189224$, where the coefficients of $u(x)$ are obtained through the third smallest eigenvalue 0.0949928 .

In order to be able to understand the discussion above more clearly, in Figures 1 and 2, we present plots of the dispersion functions of the D -optimal designs (those with $k+1$ supports are computed numerically) in certain design intervals $[a, 1]$, $-1 \leq a < 1$, for the cubic and the quartic models. There it is demonstrated how the support points of the optimal designs change according to a . It can also

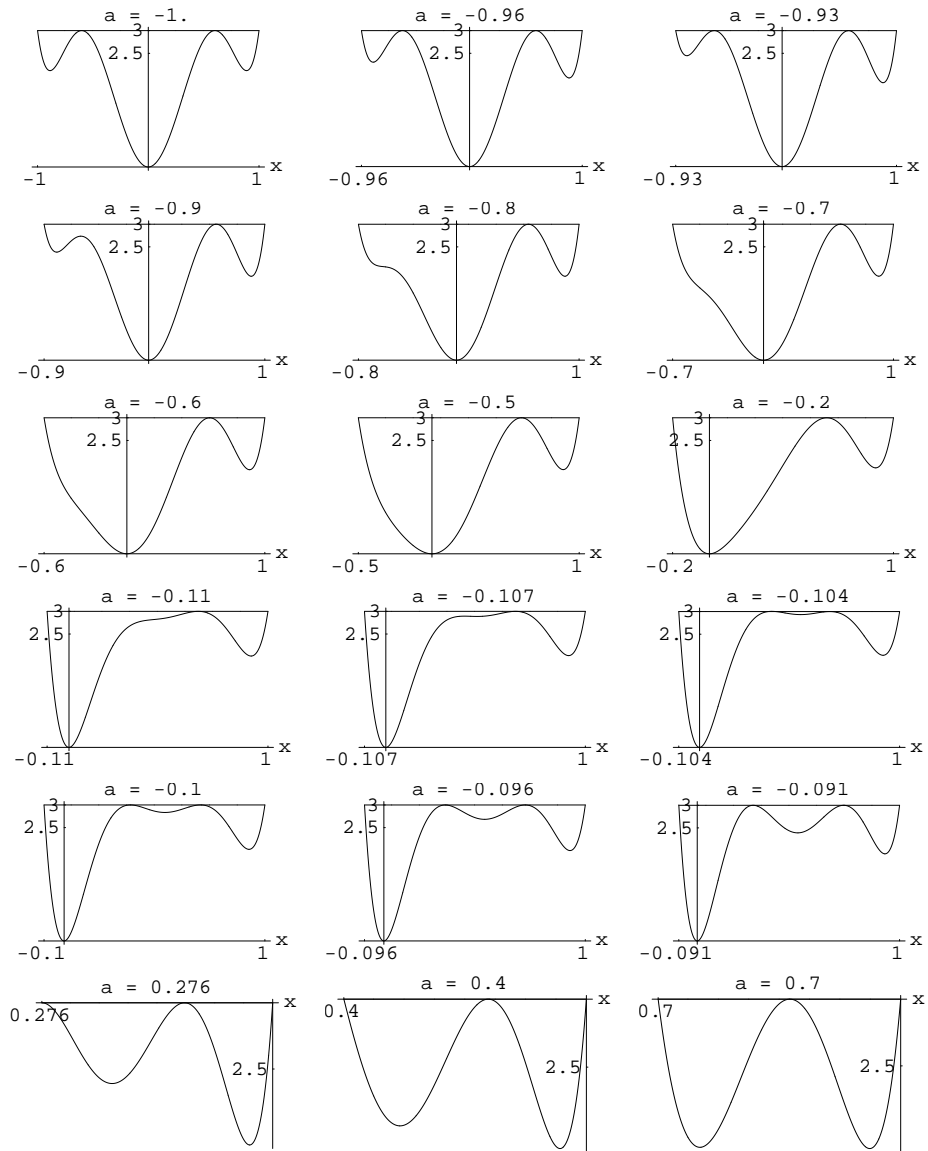


Figure 1. Plots of the dispersion functions of the D -optimal designs in certain design intervals $[a, 1]$, $-1 \leq a < 1$, for the cubic model $f_3^T(x) = (x, x^2, x^3)$.

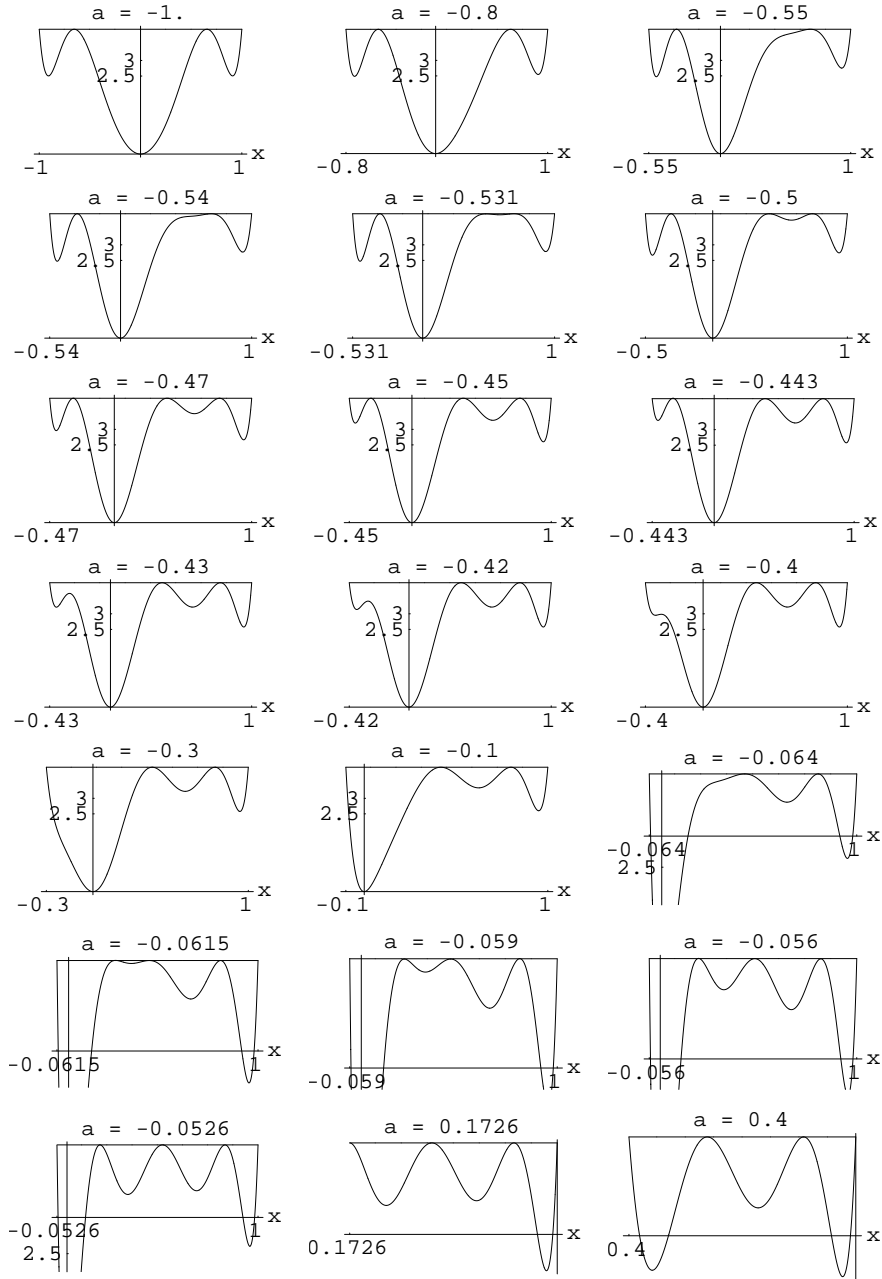


Figure 2. Plots of the dispersion functions of the D -optimal designs in certain design intervals $[a, 1]$, $-1 \leq a < 1$, for the quartic model $f_4^T(x) = (x, x^2, x^3, x^4)$.

be seen how the number of supports of the D -optimal designs change between k and $k + 1$ for different design intervals through the continuous movements of the corresponding dispersion functions.

As a remark on why we do not have similar analytic results as in Theorem 2 for the case discussed in this section, note that in Theorem 2, for $a > 0$, the Sturm-Liouville equation defined in (3.5) plays an essential role, where $\rho(x) > 0$, for $x \in (a, 1)$ is required. Although for $a < 0$, $\rho(x) < 0$ for $x < 0$ and $\rho(x) > 0$ for $x > 0$, this means the result we used concerning the Sturm-Liouville equation does not apply directly. More investigations are needed to verify the conjecture we have described in this section.

6. Approximation for $\xi_{a,k}^*$

Since the optimal designs for $f_k(x)$ are not well known, it may be of interest to compare the efficiencies of ξ_k^* for the model $f_k(x)$. For brevity, we consider only the case when $\Omega = [-1, 1]$ for $k \geq 2$, since for the case $k = 1$, ξ_k^* is still optimal. Other intervals are similarly treated. Recall that if $\bar{d}(\xi) = \max_{x \in \Omega} d(x, \xi)$, the D - and G -efficiency of a feasible design ξ are respectively given by

$$e_D(\xi) = \left(\frac{|M(\xi)|}{\max_{\eta} |M(\eta)|} \right)^{1/k} \quad \text{and} \quad e_G(\xi) = \frac{k}{\bar{d}(\xi)}.$$

To obtain efficiency bounds, let V be the Vandermonde matrix induced by the points of S_k^* , i.e. $v_{ij} = (x_j^*)^{i-1}$, and $l_i(x)$'s be the Lagrange interpolating polynomials induced by the points x_i^* , $1 \leq i, j \leq k + 1$. Clearly, $F_k(x) = V(l_1(x), \dots, l_{k+1}(x))^T = V l(x)$. Let $\mathbf{1}$ be a $(k + 1)$ vector with all elements equal to 1, $V = (\mathbf{1}, V_0^T)^T$ where V_0 is a $k \times (k + 1)$ matrix. Now $f_k(x) = V_0 l(x)$, and for any feasible design ξ ,

$$M(\xi) = \int f(x) f^T(x) d\xi(x) = \int V_0 [l(x) l^T(x)] V_0^T d\xi(x) = V_0 M_l(\xi) V_0^T,$$

where $M_l(\xi) = \int l(x) l^T(x) d\xi(x)$. After some algebra, we have

$$d(x, \xi) = l^T(x) V_0^T [L M_l^{-1} L^T - L M_l^{-1} R^T (R M_l^{-1} R^T)^- R M_l^{-1} L^T] V_0 l(x),$$

where L is a left inverse of V_0^T and $R = I - V_0^T L$. (Pukelsheim (1993, p. 74)). Let $V^{-1} = (b, L^T)$. Then $R = \mathbf{1} b^T$ and $(R M_l^{-1} R^T)^- = (b^T M_l^{-1} b)^{-1} (\mathbf{1} \mathbf{1}^T)^-$. After some computation, we have

$$\begin{aligned} & V_0^T L [M_l^{-1} - M_l^{-1} R^T (R M_l^{-1} R^T)^- R M_l^{-1}] L^T V_0 \\ &= (I - R) [M_l^{-1} - M_l^{-1} R^T (R M_l^{-1} R^T)^- R M_l^{-1}] (I - R)^T \\ &= M_l^{-1} - M_l^{-1} R^T (R M_l^{-1} R^T)^- R M_l^{-1} \\ &= M_l^{-1} - M_l^{-1} b \mathbf{1}^T (b^T M_l^{-1} b)^{-1} (\mathbf{1} \mathbf{1}^T)^- \mathbf{1} b^T M_l^{-1}; \end{aligned}$$

therefore,

$$d(x, \xi) = l^T(x)M_l^{-1}(\xi)l(x) - [l^T(x)M_l^{-1}(\xi)b]^2 / [b^T M_l^{-1}(\xi)b].$$

Since ξ_k^* is equally supported on S_k^* , $(k+1)M_l(\xi_k^*) = I$ and $l^T(x)l(x) \leq 1$ for all $x \in [-1, 1]$, we have

$$\bar{d}(\xi_k^*) = \max_x d(x, \xi_k^*) \leq k + 1, \quad (6.1)$$

and $e_G(\xi_k^*) \geq k/(k+1)$. By Theorem 2.3.4 of Kiefer (1960) and (6.1), we have $e_D(\xi_k^*) \geq [\exp(-1)]^{1/k}$ and $e_D(\xi_k^*) \leq \exp[1/(2k^2 + 2k)]^{1/k}$. Note that both the D - and G -efficiency of ξ_k^* tends to 1 as k tends to infinity.

7. Further Discussion

In this closing section, we suggest some avenues for further work in this area and a further thought of why theoretical developments for such models might be potentially important. We also discuss briefly the cases where we have submodels of $F_k(x)$ other than $f_k(x)$ and when we have more than one factor in the experiment.

We have restricted attention primarily to the case where the missing terms in $F_k(x)$ is the intercept term, but interest in other models with additional missing terms can also be useful. For example, Atkinson and Cox (1974) considered what they called the equal interest model in the context of discriminating between two plausible models. These are essentially the models considered here with the regression function having two components, x^{p-1} and x^p ; (see Hill (1974) for numerical results). It may be interesting to generalize models having more missing components on arbitrary intervals $[a, 1]$, which have been considered as a weighted polynomial regression model in Studden (1982) for interval $[0, 1]$ and Lau (1983) for interval $[-1, 1]$. For further discussion on the non-invariance properties of designs for these models under non-singular transformations, see Wong and Cook (1993).

In the work here, we have studied properties of optimal designs for the model $f_k^T(x) = (x, \dots, x^k)$ on the interval $[a, b]$, where $a < b$. By fixing $b = 1$, we have obtained analytical results when $-1/(k^2 + k - 1) \leq a \leq 1$. When a is near -1 , analytical results remain elusive and numerical results have been provided. It is interesting to note that in this paper, we did not rely on the theory of canonical moments, which is a common tool for studying D -optimal designs for the full polynomial model, (Studden (1980, 1982, 1989) and Dette (1990, 1992a) or certain reduced polynomial models, Dette (1992b)). A reason for this is we were unable to find a useful formula for the determinant of the information matrix for submodels of $F_k(x)$ of a design in terms of its canonical moments. This is unlike the case for the full model $F_k(x)$, where the determinant can be

written compactly as a finite product of functions of the canonical moments, (see Lau and Studden (1985)). We hope this work stimulates further research in constructing optimal designs for $f_k(x)$, and more generally submodels of $F_k(x)$. An immediate payoff would be the potential application to characterizing and studying properties of A -optimal designs for $F_k(x)$ using canonical moments. This is because A -optimal designs minimize $\sum_{i=1}^k |M_i(\xi)|/|M(\xi)|$, where $M(\xi)$ is the information matrix for $F_k(x)$ and $M_i(\xi)$ is the information matrix for $F_k(x)$ with the i th term deleted.

Acknowledgements

The research of the first author was supported in part by the National Science Council of the Republic of China, Grant No. NSC 83-0208-M110-003. The research of the second author was supported in part by the National Science Council of the Republic of China, Grant No. NSC 83-0208-M110-011. The authors would also like to thank the referee for many valuable comments which significantly improved this paper.

References

- Abramowitz, M. and Stegun, I. S. (1964). Handbook of Mathematical Functions. Dover, New York.
- Atkinson, A. C. and Cox, D. R. (1974). Planning experiments for discriminating between models, *J. Roy. Statist. Soc. Ser. B* **36**, 321-348.
- Birkhoff, G. and Rota, G. C. (1989). Ordinary Differential Equations, 4th edition. John Wiley, New York.
- Boyce, W. E. and DiPrima, R. C. (1992). Elementary Differential Equations and Boundary Value Problems, 5th edition. John Wiley, New York.
- Casella, G. (1983). Leverage and regression through the origin. *Amer. Statist.* **37**, 147-152.
- Detle, H. (1990). A generalization of D - and D_1 -optimal designs in polynomial regression. *Ann. Statist.* **18**, 1784-1804.
- Detle, H. (1992a). Experimental designs for a class of weighted polynomial regression models. *Comput. Statist. Data Anal.* **14**, 359-373.
- Detle, H. (1992b). Optimal designs for a class of polynomials of odd or even degree. *Ann. Statist.* **20**, 238-259.
- Fedorov, V. V. (1972). Theory of Optimal Experiments. Translated and edited by W. J. Studden and E. M. Klimko, Academic Press, New York.
- Hahn, G. J. (1977). Fitting regression models with no intercept term. *J. Quality Tech.* **9**, 56-61.
- Hawkins, D. M. (1980). A note on fitting a regression without an intercept term. *Amer. Statist.* **34**, 233-233.
- Hedayat, A., Raktoc, B. L. and Talwar, P. P. (1977). Examination and analysis of residuals: A test for detecting a monotonic relation between mean and variance in regression through the origin. *Comm. Statist. Theory Methods* **6**, 497-506.

- Hill, P. D. H. (1974). Comments on the discussion paper of Atkinson and Cox (1974). *J. Roy. Statist. Soc. Ser.B* **36**, 340-341.
- Hoel, P. G. (1958). Efficiency problems in polynomial estimation. *Ann. Math. Statist.* **29**, 1134-1145.
- Kiefer, J. (1960). Optimum experimental designs V, with applications to systematic and rotatable designs. *Proc. 4th Berkeley Symp. Math. Statist. Probab.* **1**, 381-405, Univ. of California Press, Berkeley and Los Angeles, California.
- Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems. *Canad. J. Math.* **12**, 363-366.
- Kraemer, W. (1985). The power of the Durbin-Watson test for regressions without an intercept. *J. Econometrics* **28**, 363-370.
- Lau, T. S. (1983). Theory of canonical moments and its applications in polynomial regression – part II. Technical Report No. 83-24, Purdue University, West Lafayette.
- Lau, T. S. and Studden, W. J. (1985). Optimal designs for trigonometric and polynomial regression using canonical moments. *Ann. Statist.* **13**, 383-394.
- Leoni, R. (1985). An observation about partitioning of variance in the regression model without intercept (Italian). *Rivista Di Statistica Applicata (Italy)* **18**, 233-237.
- Lim, K. S. (1992). On the stability of a threshold AR(1) without intercepts. *J. Time Ser. Anal.* **13**, 119-132.
- Perron, P. (1991). A continuous time approximation to the unstable first-order autoregressive process: The case without an intercept. *Econometrica* **59**, 211-236.
- Pukelsheim, F. (1993). *Optimal Design of Experiments*. John Wiley, New York.
- Sharma, S. C. and Ali, M. M. (1992). Robustness to non-normality of the null distribution and power of the Durbin-Watson test in regressions with and without an intercept. *J. Statist. Comput. Simulation* **42**, 93-105.
- Studden, W. J. (1980). D_s -optimal designs for polynomial regression using continued fractions. *Ann. Statist.* **8**, 1132-1141.
- Studden, W. J. (1982). Optimal designs for weighted polynomial regression using canonical moments. *Statistical Decision Theory and Related Topics III* (Edited by S. S. Gupta, J. O. Berger) **2**, 335-350, Academic Press, New York.
- Studden, W. J. (1989). Note on some Φ_p -optimal designs for polynomial regression. *Ann. Statist.* **17**, 618-623.
- Szegő, G. (1975). *Orthogonal Polynomials*, 4th edition. American Mathematical Society Colloquium Publications **23**, Providence, Rhode Island.
- Wong, W. K. and Cook, R. D. (1993). Heteroscedastic G -optimal designs. *J. Roy. Statist. Soc. Ser.B* **55**, 871-880.

Department of Applied Mathematics, National Sun Yat-sen University, Kaohsiung 80424, Taiwan.

Department of Biostatistics, University of California at Los Angeles, Los Angeles, CA 90024, U.S.A.

(Received October 1994; accepted March 1995)