ADMISSIBLE AND OPTIMAL EXACT DESIGNS FOR POLYNOMIAL REGRESSION

by

Kenneth B. Constantine[1] and Yong B. Lim[2] and W. J. Studden[3]
University of New Hampshire   Purdue University

Technical Report #85-24

Department of Statistics
Purdue University

September 1985

[1] Research supported by a NSF Graduate Fellowship, NSF Grant No. MCS-790 1707, and the University of New Hampshire Research Office.
[2] Research supported by a David Ross Grant from Purdue University.
[3] Research supported by NSF Grant No. DMS-8503771.

ADMISSIBLE AND OPTIMAL EXACT DESIGNS FOR POLYNOMIAL REGRESSION

by

Kenneth B. Constantine[1] and Yong B. Lim[2] and W. J. Studden[3]
University of New Hampshire Purdue University

ABSTRACT

The present interest is to address some classical optimal design questions in the context of the exact design setting for the polynomial regression model. A necessary condition for an exact design to be admissible is provided. It is conjectured that this condition is also sufficient and the basis for the conjecture is discussed. It is shown that the Salaevskii's conjecture holds for cubic regression. The interesting forms of G-optimal exact designs are indicated.

Key words: Polynomial Regression, Exact Design, Admissibility

AMS 1980 subject classification: 62K05, 62J05.

1. Introduction

The present study is concerned with the polynomial regression model $y(x) = \theta' f(x) + \varepsilon$, for $a \leq x \leq b$. Here $y(x)$ denotes the observed response, $\theta = (\theta_0, \theta_1, \ldots, \theta_n)'$ the (column) vector of regression coefficients, $f(x) = (1, x, \ldots, x^n)'$ the vector of regression functions, and ε the "error" random variable with mean 0 and variance σ^2. It is assumed for now that the choice of degree n is not an issue.

We suppose that N uncorrelated observations on $y(x)$ are to be obtained at "levels" $x_1, \ldots, x_N$ of x. Thus we consider the linear model $Y = X\theta + e$, where $Y = [y(x_1), \ldots, y(x_N)]'$, where $x_{ij} = f_j(x_i)$ for $1 \leq i \leq N$ and $0 \leq j \leq n$, and where $e = (\varepsilon_1, \ldots, \varepsilon_N)'$. We restrict our attention to making inferences about θ via the classical estimator $\hat{\theta} = (X'X)^{-1}X'Y$. (Here we must interpret $(X'X)^{-1}$ as a generalized inverse if rank $(X) < n+1$.) Under the preceding model's assumptions, $\hat{\theta}$ is unbiased for θ with variance/covariance matrix $V(\hat{\theta}) = \sigma^2(X'X)^{-1}$. A loose statement of the design goal, which will be made more precise, is to choose $x_1, \ldots, x_N$ so that $(X'X)^{-1}$ is "minimized".

In order to more conveniently formulate the design problem, let $x_0, \ldots, x_\ell$ now denote the distinct levels at which $n_0, \ldots, n_\ell$ observations are to be taken. (Here $n_0 + \ldots + n_\ell = N$.) An exact design ξ^N is a probability measure on $[a,b]$ which concentrates mass n_i/N at each x_i. Such a design prescribes exactly where and how to allocate observations. The set of all exact designs for a given value of N will be denoted by Ξ_N. The information matrix (per observation) of an exact design ξ^N is $M(\xi^N) = \int_a^b f(x)\, f(x)' d\xi^N(x)$. For polynomial regression, we have $m_{ij}(\xi^N) = c_{i+j}$ for $0 \leq i,j \leq n$, where each $c_k = \int_a^b x^k d\xi^N(x) = \sum_{m=0}^{\ell} n_m x_m^k / N$. It is easily seen that $V(\hat{\theta}) = \sigma^2 M^{-1}(\xi^N)/N$. Thus a reformulation of the design problem is to determine an exact design ξ^N which "minimizes" $M^{-1}(\xi^N)$.

An approach, due to Kiefer and Wolfowitz [1959], which is often taken in optimal design work is to extend consideration to the class of all approximate designs, i.e. arbitrary probability measures ξ on [a,b]. This approach has the distinct advantage of greater mathematical tractability. However, in practice, only an exact design may be implemented. It is often the case that an optimal approximate design is not exact for certain choices of N (or even for any choice of N). This limitation can be especially troublesome when N is not too large.

The present interest is to address some classical optimal design questions in the context of the exact design setting. Results will be compared with those known for approximate designs. Section 2 is devoted to the admissibility problem for polynomial regression. Theorem 2.1 provides a necessary condition for admissibility. It is conjectured that this condition is also sufficient and the basis for the conjecture is discussed. Section 3 deals with the design criterion of D-optimality. Salaevskii [1966] has conjectured that a D-optimal exact design distributes mass as equally as possible among the (n+1) support points of the D-optimal approximate design. This was shown to be true for quadratic regression by Gaffke and Kraft [1982] and shown to be false in general by Gaffke [1985]. For cubic regression it is shown that the conjecture holds. Also we summarize the cases where the conjecture is false by checking the necessary conditions for the exact D-optimal designs in Gaffke [1985]. Section 4 indicates the interesting forms of G-optimal exact designs. In particular, the well-known non-equivalence of D- and G-optimal exact designs is discussed.

2. Admissibility

Recall that the exact design problem is to determine an exact design ξ^N which "minimizes" $M^{-1}(\xi^N)$. A particular optimality criterion may correspond to a real-valued function Φ on the set of non-negative definite (n+1)x(n+1) matrices. A Φ-optimal design is one which minimizes $\Phi(M^{-1}(\xi^N))$. The cases that $\Phi(M^{-1}(\xi^N)) = |M^{-1}(\xi^N)|$ for D-optimality and $\Phi(M^{-1}(\xi^N)) = \max_{a \leq x \leq b} f(x)'M^{-1}(\xi^N)\, f(x)$ for G-optimality cases will be considered in later sections. For a concise discussion of these criteria, see Silvey [1980].

In many cases (including D- and G-optimality) the function Φ is monotone in the sense that if $M^{-1}(\xi^N) \leq M^{-1}(\tilde{\xi}^N)$, then $\Phi(M^{-1}(\xi^N)) \leq \Phi(M^{-1}(\tilde{\xi}^N))$. Here the inequality $A \leq B$ for non-negative definite matrices A and B should have the customary meaning that B-A is non-negative. Thus we are naturally led to the admissibility problem: characterize those exact designs whose inverse information matrices are minimal with respect to "$\leq$". Accordingly, an exact design ξ^N is said to be admissible if and only if there exists no other exact design $\tilde{\xi}^N$ such that $M(\xi^N) \neq M(\tilde{\xi}^N)$ and $M(\xi^N) \leq M(\tilde{\xi}^N)$. Solution of this problem will enable us to restrict our search for an optimal design to the class of admissible designs.

Note also that any inadmissible design can be excluded from consideration because it can be bettered by another design which has smaller variance for the classical estimator of any linear combination of the regression coefficients. For further motivation of the admissibility problem, see Kiefer [1959].

In the case of polynomial regression, the following lemma relates the admissibility problem to a problem involving the design moments $c_1,\ldots,c_{2n-1},c_{2n}$.

Lemma 2.1: ξ^N is admissible for polynomial regression if and only if there exists no other exact design which shares the same values of $c_1,\ldots,c_{2n-1}$ but has a larger value of c_{2n}.

Proof: See Karlin and Studden [1966].

In the approximate design setting, a design ξ is admissible for polynomial regression of degree n on [a,b] if and only if the support of ξ includes n-1 or fewer interior points. This characterization has been developed by de la Garza [1954], Kiefer [1959], and Karlin and Studden [1966]. Note that this admissibility condition involves only the support of an approximate design. It will be seen that the corresponding statement for exact designs does not involve only the support of an exact design.

The following definition establishes some terminology which will be used throughout the remainder of Section 2.

Definition 2.1: i. If $\xi^N(\{x_j\}) > 1/N$, then x_j is termed a cluster of ξ^N.
ii. If $\xi^N(\{x_j\}) = 1/N$, then x_j is a singlet of ξ^N.

Example 2.1: Let $[a,b] = [0,5]$, let $N = 10$ and let

$$\xi^N = .1\delta_0 + .1\delta_1 + .3\delta_2 + .1\delta_3 + .1\delta_4 + .3\delta_5.$$

(Here δ_x denotes a point mass at x.)

			x			x
			x			x
ξ^N:	x	x	x	x	x	x
	0	1	2	3	4	5

Thus ξ^N is comprised of an interior cluster 2, interior singlets 1,3,4, a cluster 5, and a singlet 0.

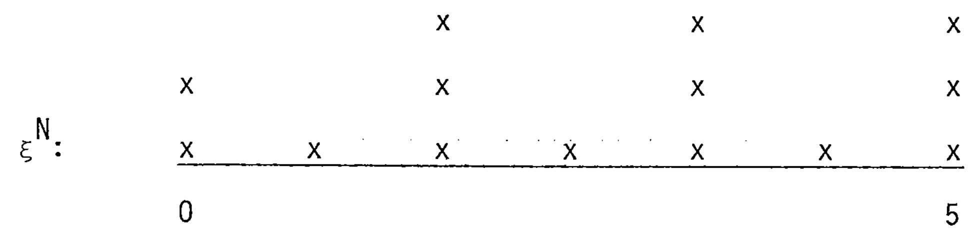
The following theorem establishes conditions that are necessary for an exact design to be admissible for polynomial regression of degree n.

Theorem 2.1: Let ξ^N be an exact design with r interior support points, with c interior clusters, and with s disjoint pairs of adjacent interior singlets. If ξ^N is admissible for polynomial regression of degree n, then

i. $r \leq 2n - 1$;

ii. $c + s \leq n - 1$.

Remark 1. Theorem 2.1 provides a complete class of exact designs for polynomial regression. A typical exact design from this class might have clusters at a and b, n-1 clusters within (a,b), and n singlets separating the clusters. For n = 3 with N = 14, this ould locate observations as follows:

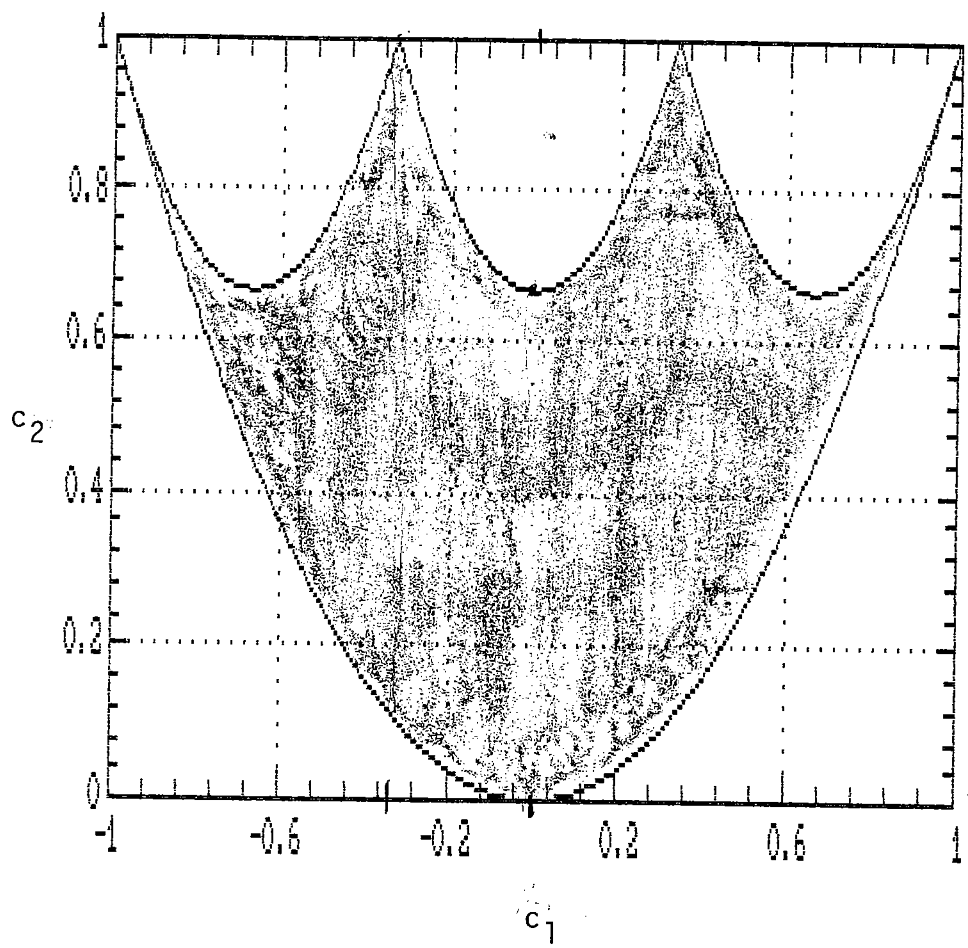


Very roughly, Theorem 2.1 turns an (N-2)-dimensional optimization problem (at least one observation must be at each end) into several (2n-1)-dimensional problems. Any subset of ξ^N satisfies the conditions of the theorem for the corresponding N and the same degree n. For example the design in example 2.1 satisifies the conditions of the theorem with N = 10 and $n \geq 3$. It is proposed that the clusters of an exact design correspond to the support points of an approximate design. The twist to this relationship is that pairs of adjacent interior singlets and singlets at a or b act as clusters.

Remark 2. At this time, the sufficiency of the conditions of Theorem 2.1 may only be conjectured. It is believed that this conjecture is valid because if an exact design satisfies i. and ii., then it is thought that no other exact design which also satisfies them can achieve the same values of $c_1, \ldots, c_{2n-1}$. If true in general, this uniqueness property (in addition to Lemma 2.1) would establish the conjecture.

In the special case of linear regression (n = 1), the validity of the preceding conjecture is readily demonstrated. Theorem 2.1 implies that an admissible exact design must have the form $\xi^N = (n_0\delta_a + \delta_n + n_2\delta_b)/N$, where $n_0 + 1 + n_2 = N$ and $a \leq x \leq b$. It is readily shown that no other exact design of this form can achieve $c_1 = (n_0 a + x + n_2 b)/N$. Figure 1 provides a sketch of moment points (c_1, c_2) when $[a,b] = [-1,1]$ and $N = 3$. The top boundary of the moment space corresponds to the admissible designs.

FIGURE 1 - Moment Space (c_1, c_2) for N=3



Proof of Theorem 2.1: The proof will make repeated use of polynomials of the form

$$P(x) = \prod_{j=1}^{2n} (x-y_j) = \sum_{m=0}^{2n} (-1)^m e_m x^{2n-m}, \tag{2.1}$$

where $a \leq y_1,\ldots,y_{2n} \leq b$, where e_k are the symmetric functions

$$e_k = \sum_{1 \leq j_1 < j_2 \ldots j_k \leq 2n} y_{j_1} y_{j_2} \cdots y_{j_k}$$

for $1 \leq k \leq 2n$, and where $e_0 = 1$. It will be convenient to define $s_k = \sum_{j=1}^{2n} y_j^k$ for $k = 1,\ldots,2n$. In terms of this notation,

$$s_k = \sum_{m=1}^{k-1} (-1)^{m+1} e_m s_{k-m} + (-1)^{k+1} k e_k$$

for $k = 1,\ldots,2n$. These equations establish a 1-1 correspondence between $s_1,\ldots,s_k$ and $e_1,\ldots e_k$ for each $k = 1,\ldots,2n$. Furthermore, it is seen that another set of points $\tilde{y}_1,\ldots,\tilde{y}_{2n}$ achieves $\tilde{s}_k = s_k$ for $k = 1,\ldots,2n-1$ but $\tilde{s}_{2n} > s_{2n}$ if and only if $\tilde{e}_k = e_k$ for $k = 1,\ldots,2n-1$ but $\tilde{e}_{2n} < e_{2n}$. According to (2.1), this is possible if and only if P can be lowered by some amount $\Delta > 0$ to the polynomial $\tilde{P}(x) = P(x) - \Delta$ with 2n roots on [a,b]. (Here $\Delta = e_{2n} - \tilde{e}_{2n}$ and the roots of $\tilde{P}$ are $\tilde{y}_1,\ldots,\tilde{y}_{2n}$.) This approach will now be applied to the admissibility problem by choosing $y_1,\ldots,y_{2n}$ appropriately.

To demonstrate that conditions i. and ii. must hold, suppose first that an exact design ξ^N has more than 2n-1 interior support points. Let $y_1,\ldots,y_{2n}$ denote any 2n of them. Then it is seen that we can lower P as in the preceding paragraph to obtain $\tilde{P}$ with associated roots $\tilde{y}_1,\ldots,\tilde{y}_{2n}$ in [a,b]. Furthermore, $\tilde{s}_k = s_k$ for $k = 1,\ldots,2n-1$ but $\tilde{s}_{2n} > s_{2n}$. Now let $\tilde{\xi}^N$ be the exact design obtained from ξ^N by exchanging a single observation at each of $y_1,\ldots,y_{2n}$ for a single observation at each of $\tilde{y}_1,\ldots,\tilde{y}_{2n}$. Then $\tilde{c}_k = c_k$ for $k = 1,\ldots,2n-1$ but $\tilde{c}_{2n} > c_{2n}$. That is, according to Lemma 2.1, ξ^N is inadmissible. Therefore, an admissible design can have no more than 2n-1 interior support points.

Suppose next that ξ^N has more than n-1 interior clusters. Let $x_1,\ldots,x_n$ denote n of them and let $y_{2i-1} = y_{2i} = x_i$ for $i = 1,\ldots,n$. As in the previous paragraph, P may be lowered to $\tilde{P}$ which has its 2n roots in [a,b]. As before, this implies that ξ^N is inadmissible and hence that no admissible design can have more than n-1 interior clusters.

To complete the proof that condition ii. is necessary, suppose that ξ^N has $c \leq n-1$ interior clusters and more than (n-1)-c disjoint pairs of adjacent interior singlets. Then let $x_1,\ldots,x_c$ denote the interior clusters, let $y_{2i-1} = y_{2i} = x_i$ for $i = 1,\ldots,c$, and let $y_{2c+1},\ldots,y_{2n}$ denote points which comprise pairs of adjacent interior singlets. By applying the same method to construct $\tilde{\xi}^N$, it is seen that ξ^N is inadmissible. Therefore, an admissible exact design with $c \leq n-1$ interior clusters can have no more than (n-1)-c disjoint pairs of adjacent interior singlets and the proof of the theorem is complete.

3. D-optimality

As already stated, a D-optimal exact design ξ_*^N minimizes $|M^{-1}(\xi^N)|$. Equivalently, $|M(\xi_*^N)| = \max_{\xi^N \in \Xi_N} |M(\xi^N)|$. An exact design ξ^N is a probability measure on [a,b] which concentrates mass $\frac{n_i}{N}$ at each x_i, $i=1,\ldots,\ell$. Without loss of generality we can assume that $[a,b] = [-1,1]$ since we have

$$|M(\xi^N)| = \sum_{1 \le i_0 < \ldots < i_n \le \ell} \frac{n_{i_0} \ldots n_{i_n}}{N^{(n+1)}} \prod_{0 \le k < j \le n} (x_{i_k} - x_{i_j})^2 \quad (3.1)$$

according to the Binet-Cauchy and Vandermonde formula.

Hoel [1958] has obtained the result that an approximate design is D-optimal for polynomial regression of degree n on $[a,b] = [-1,1]$ if and only if it concentrates equal mass at the n+1 roots of $\pi_n(x) = (1-x^2)\lambda_n'(x)$, where λ_n is the Legendre polynomial of degree n. For purposes of notation, let $-1 = x_0^* < x_1^* < \ldots < x_n^* = 1$ denote the roots of $\pi_n(x)$.

If N is an integer multiple of n+1, then the D-optimal exact design coincides with the D-optimal approximate design. Otherwise, a reasonable exact design might be one which distributes the N observations as evenly as possible among the same points $x_0^*,\ldots,x_n^*$. Salaevskii's conjecture [1966] is that such a rounded-off form characterizes the D-optimal exact designs. Salaevskii [1966] has proved that the conjecture holds for sufficiently large N and Constantine and Studden [1981] have streamlined that proof. A proof of the conjecture in the case of linear regression is not hard to provide. This can be verified directly for general N by considering figures similar to Figure 1. The value of $|M|$ is $c_2-c_1^2$ and one can check that this is maximized at the two "peaks" nearest $c_1 = 0$. For quadratic regression, Gaffke and Kraft [1982] have proven the conjecture by using the arithmetic-geometric means inequality. The elegant approach of Gaffke and Kraft [1982] was then generalized by Huang [1983] and Gaffke [1985] in order to obtain a simplified proof of the conjecture for sufficiently large N and give helpful guidelines for when N is "large enough". Their values are tabulated in Table 1.

TABLE 1

degree n	1	2	3	4	5	6	7	8	9
minimal N	2(all)	3(all)	8	15	24	35	48	63	80

For cubic regression we prove that the conjecture holds for arbitrary N in the following.

Suppose $N = k(n+1) + q$. Denote by $\xi_*^{N,T}$ a conjectured design, so that the support of $\xi_*^{N,T}$ is $\{x_0^*,\ldots,x_n^*\}$ and the weights are

$$\xi_*^{N,T}(x_i^*) = \begin{cases} \frac{k+1}{N} & \text{if } i \in T \\ \frac{k}{N} & \text{if } i \notin T, \end{cases} \qquad (3.2)$$

where $T \subset \{0,1,\ldots,n\}$ and the number of elements in T is q. By the Binet-Cauchy and Vandermonde formula we have

$$|M(\xi_*^{N,T})| = \frac{k^{(n+1-q)}(k+1)^q}{N^{n+1}} \prod_{0 \leq i < j \leq n} (x_j^* - x_i^*)^2 \quad (3.3)$$

regardless of the choice of T.

Define $g(x)' = (g_0(x),\ldots,g_n(x))$ to be the Lagrange interpolation polynomials induced by $\{x_0^*,\ldots,x_n^*\}$, i.e.,

$$g(x) = F^{-1} f(x),$$

where $F = [f(x_0^*),\ldots,f(x_n^*)]$.

Also let t_ν, $\nu = 1,\ldots,n$ be the intersection point of $g_{\nu-1}^2(x)$ and $g_\nu^2(x)$ in $[x_{\nu-1}^*, x_\nu^*]$ and then

$$g_\nu^2(x) = \max_i g_i^2(x) \ , \ x \in [t_\nu, t_{\nu+1}], \ \nu = 0,\ldots,n$$

with t_0 and t_{n+1} being equal to -1 and 1, respectively. Consider a design μ on $\{x_0^*,\ldots,x_n^*\}$ with $p_i = \mu(\{x_i^*\})$, $i = 0,\ldots,n$. Then the variance function is as follows:

$$\begin{aligned} d(x,\mu) &= f'(x)M^{-1}(\mu)f(x) \\ &= g'(x)F'[F\textstyle\int gg'd\mu \ F']^{-1}Fg(x) \\ &= \sum_{i=0}^{n} \frac{g_i^2(x)}{p_i} . \end{aligned}$$

The following first two lemmas are similar to those given by Gaffke and Kraft [1982] and by Huang [1983].

Lemma 3.1: Suppose $\sum_{i=0}^{n} g_i^2(x) \le \frac{k}{k+1} + \frac{1}{k+1} \max_i g_i^2(x)$, $x \in I$.
Then for any $x \in [t_{i_0}, t_{i_0+1}] \cap I$ and $i_0 \in T$,

$$\frac{1}{N} d(x, \xi_*^{N,T}) \le \frac{1}{k+1} .$$

Lemma 3.2: $\min_T \mathrm{Tr}[M^{-1}(\xi_*^{N,T})M(\xi^N)] \le n+1$ for any exact design ξ^N if

$$\sum_{i=0}^{n} g_i^2(x) \le \frac{k}{k+1} + \frac{1}{k+1} \max_i g_i^2(x), \text{ for all } x \in [-1,1].$$

For $n = 3$, the roots of $\pi_n(x)$ are $\{-1, -\frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}, 1\}$ and the Lagrange polynomials are obtained as follows: $g_0(x) = \frac{5}{8}(1-x)(x^2 - \frac{1}{5})$, $g_1(x) = -\frac{5\sqrt{5}}{8}(1-x^2)(x-\frac{1}{\sqrt{5}})$, $g_2(x) = g_1(-x)$ and $g_3(x) = g_0(-x)$. Simple algebra shows that $t_1 = -\frac{3(5-\sqrt{5})}{10}$ and $t_3 = \frac{3(5-\sqrt{5})}{10}$. Also it can be easily checked that

$$\sum_{i=0}^{3} g_i^2(x) \le \frac{k}{k+1} + \frac{1}{k+1} \max_i g_i^2(x), \quad k \ge 3,\ x \in [-1,1] \quad (3.4)$$

and

$$\sum_{i=0}^{3} g_i^2(x) \leq \frac{k}{k+1} + \frac{1}{k+1} \max_i g_i^2(x), \quad 1 \leq k \leq 2, \ |x| \geq .25. \quad (3.5)$$

Let $I_0 = [-1,t_1)$, $I_1 = [t_1, -.25)$, $I_2 = (.25, t_3)$, $I_3 = [t_3,1]$ and $I_4 = [-.25,.25]$. The following lemma is needed to prove that the conjecture is true for $1 \leq k \leq 2$. Proofs of Lemma 3.3 and Theorem 3.1 are in the Appendix.

Lemma 3.3: Suppose $\mu_{\vec{r}\cdot x_4}$ is a design on $\{x_0^*, x_1^*, x_2^*, x_3^*\}$ with $p_i = \dfrac{r_i + g_i^2(x_4)}{\sum_{j=0}^{3} (r_j + g_j^2(x_4))}$, where $\vec{r}' = (r_0, r_1, r_2, r_3)$, $0 \leq i \leq 3$ and $x_4 \in I_4$. Let k be 1 or 2. Then for

1 $r_j = k$, $0 \leq j \leq 3$,

2 $r_0 = k+1$, $r_j = k$, $j \neq 0$ or $r_3 = k+1$, $r_j = k$, $j \neq 3$, or

3 $r_0 = r_3 = k+1$, $r_1 = r_2 = k$,

we have

$$\sup_{x \in I_4} \prod_{j=0}^{3} (r_j + g_j^2(x_4)) < k^{4-q}(k+1)^q \quad (3.6)$$

and

$$d(x_i^*, \mu_{\vec{r}\cdot x_4}) = \max_{x \in I_i} d(x, \mu_{\vec{r}\cdot x_4}) \quad (3.7)$$

where

$$q = 1 + \sum_{j=0}^{3} r_j - 4k \quad \text{and} \quad 0 \le i \le 3.$$

Theorem 3.1: An exact N-point design ξ_*^N is D-optimal for cubic regression on [-1,1] if and only if $\xi_*^N = \xi_*^{N,T}$ for some T.

Gaffke [1985] has obtained a necessary condition for the exact D-optimal designs. The following lemma is from Gaffke [1985].

Lemma 3.4: If $\xi_*^N \in \Xi_N$ is a D-optimal exact design, then for any support point x^* of ξ_*^N, which is in the interior of [-1,1] and has weight $\geq 2/N$, it is necessary that

$$d''(x^*,\xi_*^N) \leq 0. \tag{3.8}$$

With q = 1, Gaffke [1985] has checked that condition (3.8) is violated for $n \geq 4$, $k = 1$, for $n \geq 6$, $k \leq 2$, for $n \geq 8$, $k \leq 3$. For $q = 1$ and $4 \leq n \leq 8$, numerical results were obtained by considering exact designs with an observation at each of -1 and +1 along with N-2 interior points $x_2,\ldots,x_{N-1}$, where $x_{i-2}^* < x_i < x_{i-1}^*$, $2 \leq i \leq N-1$. A system of N-2 nonlinear equations in $x_2,\ldots,x_{N-1}$ was obtained by setting the partial derivatives of $|M(\xi^N)|$ with respect to $x_2,\ldots,x_{N-1}$ equal to zero by using

TABLE 2 - Approximately D-optimal Exact Designs

for $4 \leq n \leq 8$ with $q = 1$

degree	support $\{\xi_*^N\}$	$\|M(\xi_*^N)\|$	$\|M(\xi_*^{N,T})\|$	$\left\{\frac{\|M(\xi_*^{N,T})\|}{\|M(\xi_*^N)\|}\right\}^{\frac{1}{n+1}}$
4	±1, ±.115384, ±.662934	.34772E-4	.34539E-4	.99866
5	±1, ±.772243, ±.354109, 0.	.71125E-7	.70379E-7	.99824
6	±1, ±.832819, ±.492047, ±.123890	.37090E-10	.35614E-10	.99421
7	±1, ±.873389, ±.605283, ±.273891, 0.	.47184E-14	.44841E-14	.99365
8	±1, ±.900592, ±.683696, ±.391883, ±.108250	.15062E-18	.14064E-18	.99241

$$|M(\xi^N)| = (\frac{1}{N})^{n+1} \cdot |\sum_{j\neq i} f(x_j)f(x_j)'|(1+f(x_i)'(\sum_{j\neq i} f(x_j)f(x_j)')^{-1}f(x_i)),$$

$i = 2,\ldots,N-1$. The resulting solution yielded the designs which are given in Table 2 and they seem to be approximately the D-optimal exact designs.

Now let us examine condition (3.8) for the conjectured designs $\xi_*^{N,T}$ with $2 \leq q \leq n$. Suppose $T = \{i_0,\ldots,i_{q-1}\}$. Noting $g_i(x) = \frac{\pi(x)}{(x-x_i^*)\pi'(x)}$ and $2x\lambda_n'(x) + (x^2-1)\lambda_n''(x) = n(n+1)\lambda_n(x)$, we get

$$g_i'(x) = \frac{n(n+1)(x-x_i^*)\lambda_n(x) + (1-x_i^{*2})\lambda_n'(x)}{n(n+1)\lambda_n(x_i^*)(x-x_i^*)^2}$$

and

$$g_i''(x) = \frac{n(n+1)\lambda_n'(x)(x-x_i^*)^2 - 2n(n+1)\lambda_n(x)(x-x_i^*) - 2(1-x^2)\lambda_n'(x)}{n(n+1)\lambda_n(x_i^*)(x-x_i^*)^3}$$

For $j \neq i$, $g_i'(x_j^*) = \frac{\lambda_n(x_j^*)}{(x_j^* - x_i^*)\lambda_n(x_i^*)}$ and, using L-Hospitals rule, we get

$g_i'(x_i^*) = 0$ and $g_i''(x_i^*) = -\frac{n(n+1)}{3(1-x_i^{*2})}$. Also $\sum_{i=0}^{n} g_i^2(x) = 1 - \frac{(1-x^2)}{n(n+1)}[\lambda_n'(x)]^2$

(Guest [1958]), and then $\frac{d^2}{dx^2}\sum_{i=0}^{n} g_i^2(x)\Big|_{x=x^*_{i_0}} = -2\frac{n(n+1)}{1-x^*_{i_0}}\lambda_n^2(x^*_{i_0})$. Thus

we get

$$d''(x^*_{i_0},\xi_*^{N,T}) = \frac{N}{k}\cdot\frac{d^2}{dx^2}\Big[\sum_{i=0}^{n} g_i^2(x) - \frac{1}{(1+k)}\sum_{i=0}^{q-1} g_{i_j}^2(x)\Big]\Big|_{x=x^*_{i_0}}$$

$$= \frac{N}{k}\cdot\frac{2n(n+1)}{1-x^{*2}_{i_0}}\Big[\frac{1}{3(1+k)} - \lambda_n^2(x^*_{i_0}) - \frac{1}{(1+k)n(n+1)}\sum_{j=1}^{q-1}\frac{1-x^{*2}_{i_0}}{(x^*_{i_0}-x^*_{i_j})^2}\frac{\lambda_n^2(x^*_{i_0})}{\lambda_n^2(x^*_{i_j})}\Big] \quad (3.9)$$

For $q \leq 3$, we choose $i_0 = [\frac{n}{2}]$, $i_1 = 0$ and $i_2 = n$ since we want to make $d''(x^*_{i_0},\xi_*^{N,T})$ as big as possible and $1 = \lambda_n(x^*_0) > \lambda_n(x^*_1) > \ldots > \lambda_n(x^*_{[\frac{n}{2}]})$ (Szegö [1978], p 164). Also $\lambda_n(x^*_{[\frac{n}{2}]})$ is equal to $\frac{1\cdot 3\cdots(n-1)}{2\cdot 4\cdots n}$ if n is even and less than $\lambda_{n-1}(0)$ otherwise (Szegö [1978], p 165). From the interlacing property of the roots of orthogonal polynomials (Freud [1971], p 17) and the evaluation of the roots of $\lambda_n'(x)$ up to degree 9, $x^*_{[\frac{n}{2}]}$ is in (-.5,0].

Noting $\frac{1-x^2}{(x-1)^2}$ is decreasing in (-.5,0], we get

$$d''(x^*_{[\frac{n}{2}]},\xi_*^{N,T}) > \frac{N}{k}\cdot\frac{2n(n+1)}{1-x^*_{[\frac{n}{2}]}}\cdot\Big(\frac{1}{3(1+k)} - \lambda_n^2(x^*_{[\frac{n}{2}]})(1+(3+1)\frac{1}{(1+k)n(n+1)})\Big) \quad (3.10)$$

from (3.9). Since R.H.S of (3.10) is increasing in n, $d''(x^*_{[\frac{n}{2}]}, \xi_*^{N,T})$ is positive for $n \geq 6$ with $k = 1$, $n \geq 8$ with $k = 2$ and 3 and $n \geq 10$ with $k = 4$, etc. With $k = 1$, numerical evaluations of (3.9) say that the conjecture is false for $4 \leq n \leq 9$ except $q = 4$, $n = 4$ and $q = 5$, $n = 5$. Also from (3.9) we can say that the conjecture is false for large enough n with any k and q.

4. G-Optimality

As already noted, a G-optimal exact design ξ_0^N minimizes $\max_{a \leq x \leq b} d(x,\xi^N)$, where the variance function

$$d(x,\xi^N) = f(x)'M^{-1}(\xi^N)\ f(x).$$

Guest [1958] obtained the G-optimal approximate design ξ_0 for polynomial regression of degree n on [-1,1]. This later turned out to coincide with the D-optimal design given by Hoel [1958], leading Kiefer and Wolfowitz [1960] to prove that the two criteria are equivalent in the general approximate design setting.

For polynomial regression of degree n, the G-optimal exact design coincides with the G(and D)-optimal approximate design when N is a multiple of n+1. Otherwise, G-optimal exact designs can exhibit some interesting behavior as indicated by the following examples.

Example 4.1: In the case of simple linear regression on [-1,1], it is well-known (see, for example Federov [1972]) that the G-optimal exact design for $N = 2k+1$ has k observations each at $x = -1$ and $x = +1$ as well as an interior singlet at $x = 0$.

Example 4.2: Consider now the setting of quadratic regression on $[-1,+1]$. For $N = 3k+1$, it is believed that the form of a G-optimal exact design is k observations each at $x = -1$ and $x = +1$, $k - 1$ observations at $x = 0$, and an interior singlet each at $x = -u$ and $x = +u$ $(u > 0)$. Among such designs, G-optimality will be attained if and only if $d(0,\xi_0^N) = d(1,\xi_0^N)$. Manipulation of this condition yields

$$(3-5/k)u^4 - (9-1/k)u^2 + 2 = 0 .$$

An interesting consequence of this result is that $u^2 \to (9 - \sqrt{57})/6$ as $k \to \infty$. That is, for large k, the G-optimal exact design for $N = 3k + 1$ has singlets at approximately $\pm.4916$. Perhaps even more interesting is that $u^2 = \sqrt{5} - 2$ for $k = 1$. Thus these two singlets at $\pm.4859$ are already very close to their asymptotic values. Note finally that $\max_{-1\leq x\leq 1} d(x,\xi_0^N) = 1 + [-1+(3k+1)(k+u^4)/2(k+u^2)^2]^{-1}$ whereas $\max_{-1\leq x\leq 1} d(x,\xi_0) = 3$.

For $N = 3k+2$, it is believed that the form of a G-optimal exact design is k observations each at $x = -1$, $x = 0$, and $x = +1$, and an interior singlet each at $x = -v$ and $x = +v$ $(v > 0)$. Among such designs, G-optimality will again be attained if and only if $d(0,\xi_0^N) = d(1,\xi_0^N)$. Manipulation of this condition in the present case yields

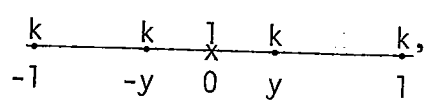
$$(3-4/k)v^4 - (9-2/k)v^2 + 4 = 0.$$

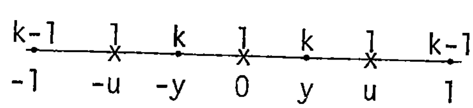
In the limit as $k \to \infty$, $v^2 \to (9-\sqrt{33})/6$ and the singlets converge to approximately $\pm.7366$. For $k = 1$, the singlets are at $\pm.7288$, already close to their limiting values. Note finally that

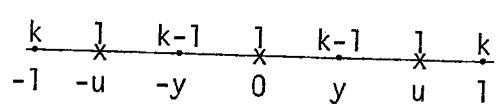
$$\max_{-1\leq x\leq 1} d(x,\xi_0^N) = 1 + [-1+(3k+2)(k+v^4)/2(k+v^2)^2]^{-1}.$$

<u>Example 4.3</u>: The final example for G-optimality is that of cubic regression on [-1,+1]. Of course when $N = 4k$, the G-optimal exact design coincides with the G (and D)-optimal approximate design. The remaining cases are now considered in turn.

For $N = 4k+1$, the following possible forms for a symmetric G-optimal exact design have been considered:

i. 

ii.  , and

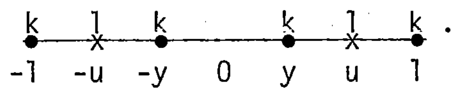
iii.  .

Note first that each type belongs to the complete class described in Section 2. For each form it was required that the variance function attain its maximum value on [-1,+1] at the points $x = \pm 1$ and $x = \pm x_0$. Thus the variance function for a cubic regression model with a design symmetric about $x = 0$ must have the form

$$d(x,\xi^N) = \alpha + \beta(x^2-1)(x^2-x_0^2)^2.$$

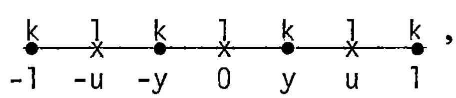
Now $d(x,\xi^N) = f(x)'M^{-1}(\xi^N)f(x)$ is a sixth degree polynomial in x with its coefficients functions of y and, in cases ii. and iii., u. Equating the coefficients of the two representation of $d(x,\xi^N)$ and eliminating all other unknowns, a constraint on y and, possibly, u is obtained. (This constraint assures that the maximum variance is attained at four points in [-1,+1] and not two.) For forms ii. and iii., a G-best design is then found by minimizing $\alpha = \max_{-1\le x\le 1} d(x,\xi^N)$ with respect to y and u by the method of Lagrange multipliers. The resultant equations were solved using the IMSL subroutine ZSCNT for N = 5,9,...,97. Interestingly, the G-best designs of types i. and ii. were, for every value of N considered, found to be inferior to the G-best designs of type iii.. The latter are presented in Table 3. For the sake of comparison, the table includes the values of 4 + 1/k which are the maximum variances achieved by distributing observations as equally as possible among the support points of the G (and D)-optimal approximate design. Of course the table shows y approaching $1/\sqrt{5}$ as N increases. Interestingly, the singlets at $\pm u$ rapidly approach asymptotic values of roughly $\pm.6935$.

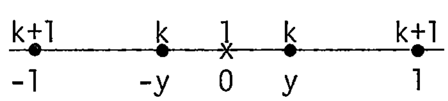
For N = 4k+2, the only form of design considered is the following:

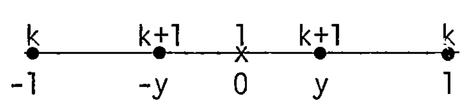


A G-best design of this form was found for N = 6,10,...,98 by the same method employed when N = 4k+1. Results appear in Table 4 along with the maximum variance 4 + 2/k obtained by rounding off the G (and D)-optimal approximate design. Of course y once again approaches $1/\sqrt{5}$ but not as rapidly as it did when N = 4k+1. Additionally the singlets at $\pm u$ rapidly approach their limiting values which, this time, are roughly $\pm.8335$.

For the final case, when N = 4k+3, the forms of design considered are the following:

i. 

ii.  , and

iii. 

For forms i. and iii., a G-best design was found for $N = 7,11,\ldots,99$ by the same methods used when $N = 4k+1$. In every instance, form i. was superior to form iii.. Surprisingly, form ii. would not allow the equalization condition on $d(x,\xi^N)$ to be met for any choice of y. Results for type i. designs are given in Table 5 and compared with the rounded-off maximum variance $4 + 3/k$. From Table 5 it is seen that y again approaches $1/\sqrt{5}$ but this time from above. The singlets at $\pm u$ again rapidly approach their limiting values which this time are roughly $\pm.8935$.

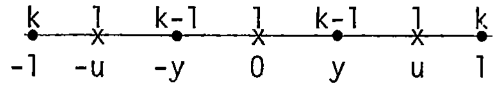
Summarizing the results for all three cases, it is seen that progressively better G-optimality performance, relative to a rounded-off design, can be achieved by the exact designs obtained here when $N = 4k+1$, $4k+2$, and $4k+3$ (respectively). Also, it is suggested that a G-optimal design for polynomial regression of any order will always have adjacent clusters separated by singlets (with no singlet at 0 for N even).

Finally, the G-optimality of designs obtained in this section would be established upon the proof of the following two conjectures.

<u>Conjecture 4.1</u>: A G-optimal exact design for polynomial regression on any (compact) interval must be symmetric about the midpoint of the interval.

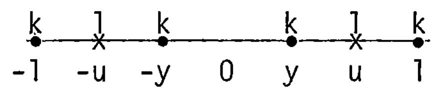
<u>Conjecture 4.2</u>: A symmetric, G-optimal exact design for polynomial regression of degree n must satisfy the equalization condition that its maximum variance is attained at exactly n+1 points.

TABLE 3 - G-optimal Designs for $n = 3$ and $N = 4k+1$ of the Form



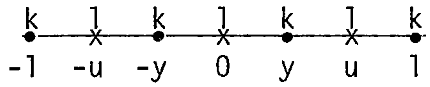
N	y	u	α	4+1/k
5	.4455	.7154	4.623	5.000
9	.4398	.6919	4.310	4.500
13	.4431	.6924	4.201	4.333
17	.4443	.6927	4.149	4.250
21	.4450	.6928	4.118	4.200
25	.4454	.6930	4.098	4.167
29	.4457	.6931	4.084	4.143
33	.4459	.6931	4.073	4.125
37	.4461	.6932	4.065	4.111
41	.4462	.6932	4.058	4.100
45	.4463	.6933	4.053	4.091
49	.4464	.6933	4.048	4.083
93	.4468	.6934	4.025	4.043
∞	.4472		4	4

TABLE 4 - G-optimal Designs for $n = 3$ and $N = 4k+2$ of the Form



N	y	u	α	4+2/k
6	.3737	.8300	4.767	6.000
10	.4036	.8309	4.368	5.000
14	.4164	.8316	4.242	4.667
18	.4235	.8320	4.180	4.500
22	.4279	.8323	4.143	4.400
26	.4309	.8325	4.119	4.333
30	.4331	.8327	4.102	4.286
34	.4348	.8327	4.089	4.250
38	.4361	.8329	4.079	4.222
42	.4372	.8330	4.071	4.200
46	.4381	.8330	4.064	4.182
50	.4388	.8331	4.059	4.167
98	.4430	.8334	4.029	4.083
∞	.4472		4	4

TABLE 5 - G-optimal Designs for $n = 3$ and $N = 4k+3$ of the Form



N	y	u	α	4+3/k
7	.4916	.8922	4.814	7.000
11	.4739	.8953	4.413	5.500
15	.4659	.8964	4.278	5.000
19	.4615	.8969	4.210	4.750
23	.4587	.8972	4.168	4.600
27	.4569	.8974	4.140	4.500
31	.4555	.8976	4.121	4.429
35	.4545	.8977	4.106	4.375
39	.4537	.8978	4.094	4.333
43	.4531	.8979	4.085	4.300
47	.4526	.8979	4.077	4.273
99	.4496	.8983	4.035	4.125
∞	.4472		4	4

Appendix

Proof of Lemma 3.3:

By the direct calculation, it can be easily checked that

$$\sup_{x\in I_4} \prod_{j=0}^{3} (r_j + g_j^2(x)) < k^{4-q}(k+1)^q, \text{ where } q = 1 + \sum_{j=0}^{3} r_j - 4k, \text{ for each case.}$$

For the second part we take $N^* = 4k+1$. By Lemma 3.1 $\xi_*^{N^*,T}$ with $T = \{i_0\}$ has local maximum at $x_{i_0}^*$, where $i_0 = 1$ or 2. So

$$[2\sum_{j=0}^{3} (r_j+g_j^2(x_4))]^{-1} d''(x_{i_0}^*, \mu_{\vec{r}\cdot x_4}) = \sum_{j\neq i_0} \frac{(g_j'(x_{i_0}^*))^2}{(r_j+g_j^2(x_4))} + \frac{g_{i_0}''(x_{i_0}^*)}{k+g_{i_0}^2(x_4)}$$

$$< \sum_{j\neq i_0} \frac{[g_j'(x_{i_0}^*)]^2}{k} + \frac{g_{i_0}''(x_{i_0}^*)}{k+1}$$

$$= (2N^*)^{-1} \cdot d''(x_{i_0}^*, \xi_*^{N^*,T})$$

$$< 0 .$$

Thus $d(x,\mu_{\vec{r}\cdot x_4})$ has local maximums at x_1^* and x_2^*. Since $d(x, \mu_{\vec{r}\cdot x_4})$ is a 6-th order polynomial, it has at most two local maximums. Hence it suffices to compare $d(x_i^*,\mu_{\vec{r}\cdot x_4})$ with $d(x,\mu_{\vec{r}\cdot x_4})$ at the boundary value of I_i, $0 \leq i \leq 3$. For I_0,

$$[\sum_{j=0}^{3}(r_j+g_j^2(x_4))]^{-1}(d(x_0^*,\mu_{\vec{r}\cdot x_4}) - d(t_1,\mu_{\vec{r}\cdot x_4})) = \frac{1}{r_0+g_0^2(x_4)} - \sum_{j=0}^{3}\frac{g_j^2(t_1)}{(r_j+g_j^2(x_4))}$$

$$> \frac{1-g_0^2(t_1)}{k+1+\sup_{x\in I_4} g_0^2(x)} - \sum_{j=1}^{3}\frac{g_j^2(t_1)}{k+\inf_{x\in I_4} g_j^2(x)}$$

$$> 0.$$

The last inequality can be shown by direct calculation. By similar arguments and simple calculation, $d(x,\mu_{\vec{r}\cdot x_4})$ is maximized at $x = x_i^*$ on I_i, $0 \le i \le 3$.

Proof of Theorem 3.1

Let ξ^N be a design in Ξ_N whose support is $\{z_1,\ldots,z_N\}$, where some of z_i's might be same. By the geometric-arithmetic means inequality, we have

$$|M(\xi^N)| \le |M(\mu)| \; Tr[\tfrac{1}{4}M^{-1}(\mu)M(\xi^N)]^4. \qquad (A.1)$$

Define

$$\Phi_N(\mu,\xi^N) = \frac{|M(\mu)|}{|M(\xi_*^{N,T})|} \; Tr[\tfrac{1}{4} M^{-1}(\mu)M(\xi^N)]^4. \qquad (A.2)$$

Then, together with (A.1) and (A.2), we get

$$|M(\xi^N)| \le |M(\xi_*^{N,T})| \cdot \Phi_N(\mu,\xi^N). \qquad (A.3)$$

So to establish the theorem, it suffices to show that for any $\xi^N \in \Xi_N$, there exists a design μ such that

$$\Phi_N(\mu,\xi^N) \leq 1. \tag{A.4}$$

Consider a design $\xi_{\vec{r}\cdot x_4}$ on $\{x_0^*,x_1^*,x_2^*,x_3^*,x_4\}$ with $\xi_{\vec{r}\cdot x_4}(x_i^*) = \frac{r_i}{N}$, $0 \leq i \leq 3$ and $\xi_{\vec{r}\cdot x_4}(x_4) = \frac{1}{N}$, where $x_4 \in I_4$, $\vec{r}' = (r_0,r_1,r_2,r_3)$ and $N = \sum_{j=0}^{3} r_j+1$ and recall $N = 4k+q$. Then,

$$\Phi_N(\mu,\xi_{\vec{r}\cdot x_4}) \leq 1 \Leftrightarrow \mathrm{Tr}[|M(\mu)|^{\frac{1}{4}} M^{-1}(\mu)M(\xi_{\vec{r}\cdot x_4})]^4 \leq |M(\xi_*^{N,T})|\ 4^4$$

$$\Leftrightarrow [(\prod_{i=0}^{3} p_i^{\frac{1}{4}})\ \frac{1}{N} \sum_{i=0}^{3} \frac{(r_i+g_i^2(x_4))}{p_i}]^4 \leq (\frac{k}{N})^{4-q} (\frac{k+1}{N})^q\ 4^4$$

$$\Leftrightarrow \sum_{i=0}^{3} (\prod_{j\neq i} p_j)(r_i+g_i^2(x_4)) - 4\ k^{\frac{4-q}{4}} (k+1)^{\frac{q}{4}} (\prod_{j=0}^{3} p_j^{\frac{3}{4}}) \leq 0$$

$$\Leftrightarrow \sum_{i=0}^{3} (\prod_{j\neq i} p_j^{\frac{3}{4}})[(\prod_{j\neq i} p_j^{\frac{1}{4}})(r_i+g_i^2(x_4))-k^{\frac{4-q}{4}} (k+1)^{\frac{q}{4}} p_i^{\frac{3}{4}}] \leq 0$$

A sufficient condition for (A.4) to hold is

$$(\prod_{j\neq i} p_j^{\frac{1}{4}})(r_i+g_i^2(x_4)) - k^{\frac{4-q}{4}} (k+1)^{\frac{q}{4}} p_i^{\frac{3}{4}} \leq 0,\ 0 \leq i \leq 3. \tag{A.5}$$

Now take

$$p_i = \frac{r_i+g_i^2(x_4)}{\sum_{j=0}^{3} (r_j+g_j^2(x_4))}, \quad 0 \le i \le 3 \qquad \text{(A.6)}$$

and call the corresponding design $\mu_{\vec{r}\cdot x_4}$. By the direct substitution of (A.6) into (A.5), we get

$$[\sum_{j=0}^{3}(r_j+g_j^2(x_4))]^{-\frac{3}{4}} (r_i+g_i^2(x_4))^{\frac{3}{4}} [\prod_{j=0}^{3}(r_j+g_j^2(x_4))^{\frac{1}{4}} - k^{\frac{4-q}{4}} (k+1)^{\frac{q}{4}}] \le 0.$$

Hence a sufficient condition for $\Phi_N(\mu_{\vec{r}\cdot x_4}, \xi_{\vec{r}\cdot x_4})$ to be less than or equal to 1 is

$$\prod_{j=0}^{3}(r_j+g_j^2(x_4)) \le k^{4-q}(k+1)^q, \; x_4 \in I_4. \qquad \text{(A.7)}$$

With $k \ge 3$ we get min $Tr[M^{-1}(\xi_*^{N,T})M(\xi^N)] \le 4$ for any exact design ξ^N by (3.4) and Lemma 3.2. Choose $\mu = \xi_*^{N,T_0}$ so that $Tr[M^{-1}(\xi_*^{N,T_0})M(\xi^N)] \le 4$. Then $\Phi_N(\xi_*^{N,T_0}, \xi^N) \le 1$. Now let us consider $q = 1$, 2 and 3 cases separately with $1 \le k \le 2$.

(i) q = 1 case.

First, suppose support $\{\xi^N\}$ has one point in I_4, say x_4, and k points in I_i, $0 \le i \le 3$. With $\vec{r} = (k, k, k, k)$ (A.7) holds from Lemma 3.3. So $\Phi_N(\mu_{\vec{r}\cdot x_4}, \xi_{\vec{r}\cdot x_4}) < 1$ and $d(x_i^*, \mu_{\vec{r}\cdot x_4}) = \max_{x \in I_i} d(x, \mu_{\vec{r}\cdot x_4})$,

$0 \leq i \leq 3$, from the second part of Lemma 3.3. Thus

$$\mathrm{Tr}[M^{-1}(\mu_{\vec{r}\cdot x_4})M(\xi^N)] = \frac{1}{N}\sum_{i=1}^{N} d(z_i,\mu_{\vec{r}\cdot x_4})$$

$$\leq \frac{1}{N}\left(k\sum_{i=0}^{3} d(x_i^*,\mu_{\vec{r}\cdot x_4}) + d(x_4,\mu_{\vec{r}\cdot x_4})\right)$$

$$= \mathrm{Tr}[M^{-1}(\mu_{\vec{r}\cdot x_4})M(\xi_{\vec{r}\cdot x_4})]$$

which implies that

$$\Phi_N(\mu_{\vec{r}\cdot x_4},\xi^N) \leq \Phi_N(\mu_{\vec{r}\cdot x_4},\xi_{\vec{r}\cdot x_4}) < 1.$$

Secondly, suppose support $\{\xi^N\}$ has more than one point in I_4 and less than or equal to k points in I_i, $0 \leq i \leq 3$. Take x_4 to be the median of support $\{\xi^N\}$. Then x_4 is in I_4. Suppose x_4 is positive. Then $g_2^2(x_4) = \max_i g_i^2(x_4)$, which implies that $d(x_2^*,\mu_{\vec{r}\cdot x_4}) = \min_i d(x_i^*,\mu_{\vec{r}\cdot x_4})$. Also $d(x_0^*,\mu_{\vec{r}\cdot x_4})$ and $d(x_3^*,\mu_{\vec{r}\cdot x_4})$ are greater than $d(x_1^*,\mu_{\vec{r}\cdot x_4})$ since $g_0^2(x)$ and $g_3^2(x)$ are less than $g_1^2(x)$, $x \in I_4$. Moreover $d(x_2^*,\mu_{\vec{r}\cdot x_4})$ is greater than $d(x_4,\mu_{\vec{r}\cdot x_4})$ since

$$\left[\sum_{i=0}^{3}(k+g_i^2(x_4))\right]^{-1}(d(x_2^*,\mu_{\vec{r}\cdot x_4}) - d(x_4,\mu_{\vec{r}\cdot x_4})) = \frac{1}{k+g_2^2(x_4)} - \sum_{i=0}^{3}\frac{g_i^2(x_4)}{k+g_i^2(x_4)}$$

$$= \frac{1-g_2^2(x_4)}{k+g_2^2(x_4)} \left[1 - \sum_{i\neq 2} \frac{(k+g_2^2(x_4))g_2^2(x_4)}{(k+g_i^2(x_4))(1-g_2^2(x_4))}\right]$$

$$> \frac{1-g_2^2(x_4)}{k+g_2^2(x_4)} \left[1-3\ \frac{k+g_2^2(x_4)}{k}\ \frac{g_2^2(x_4)}{1-g_2^2(x_4)}\right]$$

$$> 0 .$$

Hence $Tr[M^{-1}(\mu_{\vec{r}\cdot x_4})M(\xi^N)] < Tr[M^{-1}(\xi_{\vec{r}\cdot x_4})M(\xi_{\vec{r}\cdot x_4})]$, which implies that $\Phi_N(\mu_{\vec{r}\cdot x_4},\xi^N) < \Phi(\mu_{\vec{r}\cdot x_4},\xi_{\vec{r}\cdot x_4}) < 1$. By the same arguments $\Phi(\mu_{\vec{r}\cdot x_4},\xi^N)$ is less than 1 for nonpositive x_4.

Thirdly, suppose support $\{\xi^N\}$ has more than k points in some I_{i_0}. Taking $\mu_{\vec{r}\cdot x_4} = \xi_*^{N,T}$ with $T = \{i_0\}$, we get

$$\Phi(\xi_*^{N,T},\xi^N) = Tr[\tfrac{1}{4}M^{-1}(\xi_*^{N,T})M(\xi^N)]$$

$$= \frac{1}{4} \cdot \frac{1}{N} \sum_{i=1}^{N} d(z_i,\xi_*^{N,T})$$

$$\leq \frac{1}{4} \cdot \frac{1}{N}\left(\frac{N}{k+1}\ (k+1) + \frac{N}{k} \cdot 3k\right)$$

$$= 1 .$$

The third inequality follows from Lemma 3.1.

(ii) $q = 2$ case.

First, suppose support $\{\xi^N\}$ has one point in I_4, say x_4, $(k+1)$ points in I_0 or I_3 and k points in all the other I_j's. With r_j, $0 \leq j \leq 3$, being the number of points in I_j, we get

$$\Phi(\mu_{\vec{r}\cdot x_4}, \xi_{\vec{r}\cdot x_4}) < 1$$

by Lemma 3.3. By the second part of Lemma 3.3,

$$Tr[M^{-1}(\mu_{\vec{r}\cdot x_4})M(\xi^N)] < Tr[M^{-1}(\mu_{\vec{r}\cdot x_4})M(\xi_{\vec{r}\cdot x_4})].$$

Hence $\Phi(\mu_{\vec{r}\cdot x_4}, \xi^N) < 1$.

Secondly, suppose support $\{\xi^N\}$ has $(2k+1)$ points in $I_0 \cup I_3$, at most k points in I_1 and I_2, and $(2k+1)$ points in $I_1 \cup I_2 \cup I_4$. If the number of points in I_0 is greater than k, set $r_0 = k+1$ and $r_j = k$, $1 \leq j \leq 3$. Otherwise, set $r_3 = k+1$ and $r_j = k$, $0 \leq j \leq 2$. Take x_4 to be the median of the points in $I_1 \cup I_2 \cup I_4$. Then x_4 is in I_4. By arguments similar to those in the $q = 1$ case, it can be easily shown that

$$\Phi(\mu_{\vec{r}\cdot x_4}, \xi^N) < \Phi(\mu_{\vec{r}\cdot x_4}, \xi_{\vec{r}\cdot x_4}) < 1.$$

Thirdly, suppose support $\{\xi^N\}$ has at least $(2k+2)$ points in $I_1 \cup I_2 \cup I_4$ or $(k+1)$ points in at least two of I_j's, $0 \leq j \leq 3$, say i_0 and i_1. Take $\mu_{\vec{r}\cdot x_4}$ to be $\xi_*^{N,T}$ with $T = \{1,2\}$ or $T = \{i_0, i_1\}$. Then

$$\Phi(\xi_*^{N,T},\xi^N) \le 1$$

follows from the Lemma 3.1.

(iii) q = 3 case

First, suppose support $\{\xi^N\}$ has one point in I_4, say x_4, (k+1) points in I_0 and I_3, and k points in I_1 and I_2. With $\vec{r}$ = (k+1, k, k, k+1), it can be easily shown that

$$\Phi(\mu_{\vec{r}\cdot x_4},\xi^N) \le \Phi(\mu_{\vec{r}\cdot x_4},\xi_{\vec{r}\cdot x_4})$$
$$< 1$$

by using Lemma 3.3.

Secondly, suppose support $\{\xi^N\}$ has (k+1) points in I_0 and I_3, at most k points in I_1 and I_2, and 2k+1 points in $I_1 \cup I_2 \cup I_4$. Take x_4 to be the median of points in $I_1 \cup I_2 \cup I_4$. Then x_4 is in I_4. Similarly to the q = 1 case, we get

$$\Phi(\mu_{\vec{r}\cdot x_4},\xi^N) \le \Phi(\mu_{\vec{r}\cdot x_4},\mu_{\vec{r}\cdot x_4}) < 1.$$

Thirdly, suppose support $\{\xi^N\}$ has k+1 points in three of I_j's, say $\{i_0,i_1,i_2\}$, or k points in I_0 or I_3. Take T = $\{i_0,i_1,i_2\}$ or T = {1,2,3} or T = {0.1,2}. Then

$$\Phi(\xi_*^{N,T},\xi^N) \le 1$$

follows from Lemma 3.1.

For the only if part, we assume that ξ_*^N is D-optimal. Then $|M(\xi_*^N)| = |M(\xi_*^{N,T})|$. So the equality holds in the geometric-arithmetic means inequality, which implies that $M(\xi_*^N)$ is equal to $M(\xi_*^{N,T})$ for some T. By the Theorem 2.1 in Karlin and Studden (1966) $M(\xi_*^{N,T})$ is a boundary point of $\{M(\xi): \ \xi$ is a measure on $[-1,1]\}$ and $\xi_*^{N,T}$ has the unique representation. Thus $\xi_*^N = \xi_*^{N,T}$ for some T.

REFERENCES

Constantine, K. B. and Studden, W. J. [1981]. Optimal Exact Designs for Polynomial Regression. Purdue University Mimeograph Series #81-35.

Federov, V. V. [1972]. Theory of Optimal Experiments. Translated by Studden, W. J. and Klimko, E. M., Academic Press, New York.

de la Garza, A. [1954]. Spacing of information in polynomial regression. Ann. Math. Statist., Vol. 25, 123-130.

Freud, G. [1972]. Orthogonal Polynomials. Pergamon Press, New York.

Gaffke, N. and Kraft, O. [1982]. Exact D-optimum designs for quadratic regression. J. Roy. Statist. Soc. Ser. B, Vol. 44, 394-397.

Gaffke, N. [1985]. On D-optimality of exact linear regression designs with minimum support. (to appear in J.S.P.I.)

Guest, P. G. [1958]. The spacing of observations in polynomial regression. Ann. Math. Statist., Vol. 29, 294-299.

Hoel, P. G. [1958]. Efficiency problems in polynomial estimation. Ann. Math. Statist., Vol. 29, 1134-1145.

Huang, M. L. [1983]. Design Problems in Model Robust Regression and Exact D-optimality. Purdue University Mimeograph Series #83-32.

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

Kiefer, J. [1959]. Optimum experimental designs. J. Roy. Statist. Soc. Ser. B, Vol. 21, 273-319.

Kiefer, J. and Wolfowitz, J. [1959]. Optimum Designs in Regression Problems. Ann. Math. Statist., Vol. 30, 271-294.

Kiefer, J. and Wolfowitz, J. [1960]. The equivalence to two extremum problems. Can. J. Math., Vol. 12, 363-366.

Salaevskii, O. V. [1966]. The problem of the distribution of observations in polynomial regression. Proc. Steklov Inst. Math. #79, 147-166.

Silvey, S. D. [1980]. Optimal Design. Chapman and Hall, New York.