Efficient D_s-optimal designs for multivariate polynomial regression on the q-cube

by

Yong B. Lim* and W. J. Studden
Louisiana State University | Purdue University

Technical Report 86–49

Department of Statistics
Purdue University

Revised June 1987

* Research supported by NSF Grant No. DMS–8503771.

Efficient D_s-optimal designs for multivariate polynomial regression on the q-cube

by

Yong B. Lim* and W. J. Studden
Louisiana State University — Purdue University

Abstract

Polynomial regression in q variables, of degree n, on the q cube is considered. D_s-optimal designs for estimating higher degree terms are investigated. Numerical results are given for $n = 2$ with arbitrary q and for $n = 3, 4, 5$ and $q = 2, 3$. Exact solutions are given within the class of product designs together with some efficiency calculations.

* Research supported by NSF Grant No. DMS–8503771.

1. Introduction

Consider the polynomial regression model in q variables, of degree n, on the q-cube. Thus it is assumed that for each $x = (x_1, \ldots, x_q)$ in the q-cube

$$\mathcal{X} = \{x : | x_i | \leq 1, \quad i = 1, \ldots, q\} \tag{1.1}$$

a random variable $Y(x)$ with mean $\Sigma_{i=1}^{K} \theta_i f_i(x) = f(x)'\theta$ and variance σ^2, independent of x, can be observed. Here the regression functions $f_i(x)$ are known functions of the form $\Pi_{j=1}^{q} x_j^{m_j}$ where m_j are nonnegative integer with sum $\leq n$. It is well known (e.g. Scheffe (1958)) that the number of such functions is $\binom{n+q}{q}$.

A design ξ is a probability measure on $\mathcal{X}$. The information matrix is given by

$$M(\xi) = \int f(x)\, f(x)'\, \xi(dx). \tag{1.2}$$

If the design is implementable and N uncorrelated observations are taken, then the covariance matrix of the least squares estimates $\hat{\theta}$ of θ is given by

$$\text{Var}\,(\hat{\theta}) = \frac{\sigma^2}{N} M^{-1}(\xi). \tag{1.3}$$

Much of the "Kiefer type" optimal design theory is concerned in minimizing some functional of $M^{-1}(\xi)$ over ξ.

The basic criterion of design optimality we shall use here is that of D-optimality (or D_s-optimality) developed largely by Kiefer (1959, 1961) and Kiefer and Wolfowitz (1959, 1960). The D-optimality criterion is known, by the celebrated Kiefer-Wolfowitz theorem, to be equivalent to the G-optimality criterion. So, the design ξ^* is D-optimal iff the variance function $d(x, \xi^*) \leq K$ for all $x \in \mathcal{X}$, where $d(x, \xi^*) = f(x)'\, M^{-1}(\xi^*)\, f(x)$.

In the case where interest is in only s of the K parameters in θ, it is customary to decompose f into $f' = (f_1', f_2')$ where f_2 corresponds to the s parameters of interest. Similarly the information matrix is decomposed into

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

The covariance matrix of the s parameters is proportional to the inverse of

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

(Here we must interpret M_{11}^{-1} as a generalized inverse if rank $(M_{11}) < K - s$.) Corresponding to the D_s-optimality criterion, we have the following theorem (Kiefer (1961), Karlin & Studden (1966b) and Atwood (1969)).

Theorem 1.1 *If $M(\xi^*)$ is nonsingular, then the following assertions:*

(1) the design ξ^* maximizes $|\Sigma_s(\xi)|$

(2) the design ξ^* minimizes $\max_x d_s(x, \xi)$, where $d_s(x, \xi) = d(x, \xi) - f_1(x)'\, M_{11}^{-1}(\xi)\, f_1(x)$

(3) $\max_x d_s(x, \xi^*) = s$
are equivalent.

To find the maximum of $| \Sigma_s(\xi) |$, we use the result that

$$| \Sigma_s(\xi) | = \frac{| M(\xi) |}{| M_{11}(\xi) |}. \tag{1.4}$$

The model under consideration is invariant or symmetric with respect to the group consisting of permutations and sign changes of the coordinates. The invariance theorem (Kiefer (1959), (1961) and Giovagnoli, Pukelsheim and Wynn (1987)) which concludes that there exists a symmetric D and D_s-optimal design is a very important tool for obtaining D and D_s-optimal designs either theoretically or numerically. All of the designs we consider will be symmetric with respect to the above group.

An outline of this paper is as follows. In Section 2 we discuss the case $n = 2$. Kiefer (1961a), Kono (1962) and Farrel, et al (1967) give a complete description of a symmetric D-optimal design. We give a similar analysis for estimating only the quadratic terms and simplify the corresponding geometrical considerations. In Section 3 we give some numerical results for $q = 2, 3$ and $n = 3, 4, 5$. These results support the general idea that the D and D_s-optimal design are "close to" product designs. Thus for the cubic regression in one dimension we use 4 support points in our design while for $q = 2$ the D-optimal design is on a nearly rectangular grid of 16 point. This motivated the use of product designs in Section 5. Through the use of certain canonical moments we are able to describe more or less explicitly the D and D_s-optimal product design. These turn out to be fairly efficient. Section 4 has some preliminary discussion and lemmas regarding the canonical moments.

2. Quadratic D_s-optimal Design.

Kiefer (1961a), Kono (1962) and Farrell, et al (1967) give a rather complete description of the D-optimal design when $n = 2$ and q is arbitrary. Further considerations of a similar nature are included in Lim, Studden and Wynn (1986) where an example involving a factorial model of type $3^q 2^r$ is analyzed. Here we describe the details for estimating all of the quadratic terms. The analysis used here originates with Kiefer (1961). Our analysis of the resulting geometrical considerations are somewhat simpler.

The regression vector is $f'(x) = (f_1'(x), f_2'(x))$, where $f_1'(x) = (1, x_1, \ldots, x_q)$ and $f_2'(x) = (x_1^2, \ldots, x_q^2, x_1x_2, \ldots x_{q-1}x_q)$. The vector $f_2(x)$ contains all the quadratic terms.

By the invariance theorem, there exist D and D_s-optimal designs which are symmetric with respect to permutations and sign changes of x_i's, $i = 1, \ldots, q$. Both $d(x, \xi)$ and $d_s(x, \xi)$ are quartic functions when considered as functions of each variable separately. Moreover they are symmetric with positive coefficient for x_i^4, so that their maximum can occur only at $x_i = \pm 1$ or 0. Thus the symmetric optimal design must be supported on E, where $E = \{x : | x_i | = 0 \text{ or } 1\}$.

For symmetric designs supported on E, we let

$$u = \int x_1^2 \xi(dx) = \int x_1^4 \xi(dx) \text{ and } v = \int x_1^2 x_2^2 \xi(ds). \tag{2.1}$$

It is then easy to show (see Kiefer (1961a)) that

$$| \Sigma_s(\xi) | = \frac{| M(\xi) |}{| M_{11}(\xi) |} = v^{\frac{q(q-1)}{2}}(u-v)^{q-1}\big(u+(q-1)v-qu^2\big). \tag{2.2}$$

Some algebra shows that $| \Sigma_s(\xi) |$ is maximized at

$$u^* = \frac{(2q^2+q+5)+(q-1)\sqrt{4q^2+4q+9}}{4(q^2+q+2)} \tag{2.3}$$

$$\text{and} \quad v^* = \frac{(2q^2-q+3)u^*-(q+1)}{2q^2-2} \tag{2.4}$$

For $i = 1, 2, \ldots, q$, let E_i be the subset of E consisting of those $\binom{q}{i} \cdot 2^i$ elements with $q-i$ components of x being equal to zero. The following theorem characterizes those sets of the form $\cup_{i=1}^{3} E_{r_i}$ which can support a symmetric D_s-optimal design.

Theorem 2.1 *The set $\cup_{i=1}^{3} E_{r_i}$ supports a quadratic D_s-optimal design for quadratic regression on the q-cube if and only if*

$$0 \le r_1 \le (q-1) \cdot \frac{u^*-v^*}{1-u^*} \le r_2 \le q-1, \quad r_3 = q. \tag{2.5}$$

Proof. *For ξ supported on E the space of possible (u,v) is the convex hull of $\{(\frac{i}{q}, \frac{i(i-1)}{q(q-1)}),\ \ i=1,\ldots,q\}$ since $u = \sum_{i=0}^{q} \frac{i}{q} \cdot \xi(E_i)$* and $v = \sum_{i=0}^{q} \frac{i(i-1)}{q(q-1)} \xi(E_i)$. It is not clear at this point that u^* and v^* in (2.3) and (2.4) are of this form.

Let $z_1 = \frac{i}{q}$ and $z_2 = \frac{i(i-1)}{q(q-1)}$. Then

$$z_2 = \frac{q}{q-1} z_1^2 - \frac{1}{q-1} z_1. \tag{2.6}$$

Consider z_1 as a random variable on [0,1] and let c_1 and c_2 be the 1st and 2nd moments of z_1. The set of all possible values or moment space of $(c_1, \frac{q}{q-1}c_2 - \frac{1}{q-1}c_1)$ is the convex hull of (z_1, z_2) where $z_2 = \frac{q}{q-1}z_1^2 - \frac{1}{q-1}z_1$. Note that $z_2 = 1$ for $z_1 = 1$ and $z_2 = 0$ for $z_1 = 0$ or $1/q$. The possible values for (u,v) are a subset of these corresponding to z_1 having mass only on $z_1 = i/q, \quad i = 0, 1, \ldots, q$. Figure 2.1 provides a sketch of both sets when $q = 4$.

Figure 2.1 goes here.

Since $u^* > (q-1)/q$, we have to choose r_3 to be q, which corresponds to (1,1) in the moment space. Let L be the line which passes through (1,1) and (u^*, v^*) and u_0 be the abscissa of the intersection point of L and the lower boundary of the moment space $(c_1, \frac{q}{q-1}c_2 - \frac{1}{q-1}c_1)$. Then

$$\frac{r_1}{q} \le u_0 \le \frac{r_2}{q} \tag{2.7}$$

iff (u^*, v^*) is in the convex hull of $\{(\frac{r_1}{q}, \frac{r_1(r_1-1)}{q(q-1)}), (\frac{r_2}{q}, \frac{r_2(r_2-1)}{q(q-1)}), (1,1)\}$. Thus there exists a symmetric D_s-optimal design on $E_{r_1} \cup E_{r_2} \cup E_q$. It can be easily checked that

$$u_0 = \frac{q-1}{q}\frac{u^* - v^*}{1-u^*}. \tag{2.8}$$

By substitution of (2.8) into (2.7), we get

$$0 \le r_1 \le (q-1)\frac{u^* - v^*}{1-u^*} \le r_2, \quad r_3 \equiv q.$$

Substitute (2.3) and (2.4) into $\frac{u^*-v^*}{1-u^*}$ and use $\sqrt{4q^2+4q+q} < 2q+1+4/(2q+1)$. Then it follows that

$$\frac{u^* - v^*}{1-u^*} < 1.$$

Thus $r_2 \le q-1$, which assures the existence of a symmetric D_s-optimal design on $E_0 \cup E_{q-1} \cup E_q$.

The weights for a symmetric D_s-optimal design with $r_1 = 0$ and $r_2 = q-1$ are listed in Table 2.1 for $2 \le q \le 5$. These would be beneficial if fewer points in the design are desired. For comparison purposes, included are the weights for a symmetric D-optimal design from Kono (1962) and Kiefer (1961a). For estimating all of the quadratic terms only, more weight is on the center and E_{q-1} and less weight on the corners of the q-cube.

Table 2.1 Weights for a symmetric D- and D_s-optimal design on E with a minimal support for quadratic polynomial regression on the q-cube.

		D-optimal design	D_s-optimal design
q=2	$\xi^*(E_2)$	.538	.191
	$\xi^*(E_1)$	.321	.539
	$\xi^*(E_0)$	.096	.270
q=3	$\xi^*(E_3)$	.510	.239
	$\xi^*(E_2)$	.424	.528
	$\xi^*(E_0)$	.066	.233
q=4	$\xi^*(E_4)$	.451	.241
	$\xi^*(E_3)$	.502	.577
	$\xi^*(E_0)$	.047	.182
q=5	$\xi^*(E_5)$	.402	.230
	$\xi^*(E_4)$	.562	.625
	$\xi^*(E_0)$	.036	.145

3. Numerical D_s-optimal designs.

In this section we consider some numerical results for $q = 2,3$ and $n = 3,4,5$. For convenience we shall call ξ^* a 'numerical' D_s-optimal design if $\sup_x d_s(x, \xi^*)$ is found to be $\leq s$ to five significant digits. The five digits is somewhat arbitrary. The results were obtained on a CDC 6500 using single precision.

For $q = 2$ and $n = 3$ Farrell, *et al* (1967) considered a symmetric design ξ which put mass $w_1/4$ at $(\pm 1, \pm 1)$, $w_2/8$ at $(\pm 1, \pm a)$ and $(\pm a, \pm 1)$ and the remaining $(1 - w_1 - w_2)/4$ at $(\pm b, \pm b)$ and showed numerically that $| M(\xi) |$ was maximized at

$$a = .3588, \quad b = .4800, \quad w_1 = .3677 \quad \text{and } w_2 = .4610.$$

For this design ξ^*, they also computed $\sup_x d(x, \xi^*)$ numerically and found $\sup_x d(x, \xi^*)$ to be ≤ 10 to five decimal places.

We have done a similar analysis for the 4-th and 5-th degree regression on the 2-cube. Resulting 'numerical' symmetric D-optimal designs are listed in Table 3.1. We include the cubic case for completeness and comparison with Table 3.2. In Table 3.1 a typical point is indicated. The full design is obtained by taking permutations and sign changes of typical points. The divisors in the weight column are the number of symmetric points.

In each case we considered a 'perturbed' symmetric product design. For example with $q = 2$ and $n = 4$ we use a design with a set of typical points $\{(1,1),\ (1,a),\ (1,0),\ (b,b)\ (c,0),\ (0,0)\}$. For $n = 5$, we use $\{(1,1),\ (1,a),\ (1,b)\ (c,c),\ (d,d),\ (e,f)\ \}$. In each case the symmetry allows us to block the information matrix according to the parity of the power of each component. For $q = 2$ we divide f into 4 groups while we get 8 groups for $q = 3$. For $q = 2$ and $n = 4$

$$f'_{(1)} = (1, x_1^2, x_2^2, x_1^2 x_2^2, x_1^4, x_2^4), \qquad f'_{(2)} = (x_1, x_1 x_2^2, x_1^3),$$
$$f'_{(3)} = (x_2, x_2 x_1^2, x_2^3) \qquad \text{and} \qquad f'_{(4)} = (x_1 x_2, x_1 x_2^3, x_1^3 x_2),$$

The determinant in each case was maximized on the CDC 6500, by using the Newton-Raphson Algorithm, as a function of the 8 or 10 parameters involved, which gave the design ξ^* in Table 3.1. $Sup_x d(x, \xi^*)$ was computed numerically and found to be ≤ 15 or 21 to five decimal places. As n increases, numerical problems increase dramatically. For $n = 5$ an initial starting design was even hard to obtain. For this a program *ACED* (Algorithms for the Construction of Experimental Designs by W. J. Welch, see *American Statistician* 1985, p. 146) was used to distribute 30 observations on a grid of 681 candidate points on $[-1,1] \times [-1,1]$. For $n = 6$ the ACED seemed to give a reasonable starting design. However, based on this starting design, our optimization algorithm failed to produce an optimal design.

Similarly, we get numerical D_s-optimal designs for the highest order terms for $n = 3,4,5$ and $q = 2$ and those are listed in Table 3.2. Comparing Table 3.2 with Table 3.1, we note that the design points inside the 2-cube move toward the 4 corner points and the weights ‘shift’ toward the inside design points.

Table 3.1 Numerical symmetric D-optimal designs on the 2-cube.

	Design Point	Weight
$n = 3$	(1,1)	.3677/4
	(1,.3588)	.4610/8
	(.4800,.4800)	.1713/4
$n = 4$	(1,1)	.2473/4
	(1,.5811)	.3508/8
	(1,0)	.1582/4
	(.6442,.6442)	.1203/4
	(.6854,0)	.0722/4
	(0,0)	.0512/1
$n = 5$	(1,1)	.1785/4
	(1,.7039)	.2590/8
	(1,.2549)	.2453/8
	(.7574,.7574)	.0939/4
	(.3208,.3208)	.1079/4
	(.7446,.1963)	.1154/8

Table 3.2 Numerical symmetric D_s-optimal designs for the highest order coefficients on the 2-cube.

	Design Point	Weight
$n = 3$	(1,1)	.2606/4
	(1,.3680)	.4665/8
	(.5207,.5207)	.2729/4
$n = 4$	(1,1)	.1596/4
	(1,.6170)	.3382/8
	(1,0)	.1516/4
	(.6876,.6876)	.1814/4
	(.7453,0)	.0891/4
	(0,0)	.0801/1
$n = 5$	(1,1)	.1089/4
	(1,.7336)	.2263/8
	(1,.2775)	.2265/8
	(.7951,.7951)	.1406/4
	(.3393,.3393)	.1521/4
	(.7829,.1922)	.1456/8

Two further cases were considered. The D-optimal design for $n = 3$ and $q = 3$ and D_s-optimal design for the two highest order coefficients for $n = 3$ and $q = 2$ are given in

Table 3.3 and Table 3.4.

Table 3.3 Numerical symmetric D-optimal design for $n = 3$ and $q = 3$.

Design Point	Weight
(1,1,1)	.3142/8
(1,1,.2970)	.3942/24
(1,.4215, .4215)	.2649/24
(.5012, .5012, .5012)	.0267/8

Table 3.4 Numerical symmetric D-optimal design for the two highest order coefficients for $n = 3$ and $q = 2$.

Design Point	Weight
(1,1)	.3241/4
(1,.3360)	.4490/8
(.4579,.4579)	.2269/4

We remark that a symmetric numerical D-optimal design ξ^* for $q = 2$ and $n = 3$ is unique. As in Farrel, et. al (1967), this can be shown by checking that $\{x : d(x, \xi^*) - 10 = 0\}$ is exactly the support of ξ^* and a 27×16 matrix $\|\phi_i(\mathbf{x}_j)\|$, where $\phi_i(x)$ is of the form $\Pi_{j=1}^{q} x_j^{m_j}$ with $1 \le \sum_{j=1}^{q} m_j \le 6, m_j \ge 0$ and $\mathbf{x}_j \in$ support ξ^*, has full column rank. But for $n = 4$ and 5, the D-optimal design may not be unique since the matrix $\|\phi_i(\mathbf{x}_j)\|$ does not have full column rank.

4. Canonical Moments

In this section we describe some results concerning canonical moments used in the next section.

For an arbitrary measure ξ on [-1,1] let $c_k = \int_{-1}^{1} x^k d\xi(x)$. For a given finite set of moments $c_0, \ldots, c_{i-1}$, let c_i^+ denote the maximum of the i-th moment $\int_{-1}^{1} x^i d\xi(x)$ over the set of all measures ξ having the given set of moments $c_0, \ldots, c_{i-1}$. Similarly let c_i^- denote the corresponding minimum. The canonical moments are defined by

$$p_i = \frac{c_i - c_i^-}{c_i^+ - c_i^-}, \qquad i = 1, 2, \ldots \tag{4.1}$$

The canonical moments p_i range freely over [0,1] and permit easy maximization of $| M(\xi) |$ when $q = 1$. The remainder of the problem is converting the optimum p_i either to the support points and the weights in the corresponding design. Most of the proofs of the following lemmas are in either Studden (1982) or Lau (1983).

Lemma 4.1 *The design ξ is symmetric iff $p_{2i+1} = \frac{1}{2}$ for all i.*

Let $p_0 = 0$ and define $q_i = 1 - p_i, \quad i \geq 0$. Also define

$$\varsigma_i = q_{i-1} p_i, \qquad i = 1, 2, \ldots \tag{4.2}$$

Let a sequence of polynomials $W_\ell(x), \ell \geq 0$ be defined by taking them orthogonal to $d\xi$. Then the recursive relation for the orthogonal polynomials $W_\ell(x)$ with leading coefficient 1 is given as follows:

Lemma 4.2 *Let $W_0(x) = 1$ and $W_1(x) = x$. Then the orthogonal polynomials $W_\ell(x)$ for $\ell \geq 2$ satisfy the recursive relations*

$$W_\ell(x) = (x + 1 - 2\varsigma_{2\ell-2} - 2\varsigma_{2\ell-1}) W_{\ell-1}(x) - 4\varsigma_{2\ell-3}\varsigma_{2\ell-2} W_{\ell-2}(x). \tag{4.3}$$

The following lemma expresses the L_2 norm of an orthogonal polynomial $W_\ell(x)$ in terms of the canonical moments.

Lemma 4.3 *For $\ell \geq 1$,*

$$\int_{-1}^{1} W_\ell^2(x) dx = 2^{2\ell} \cdot \varsigma_1 \varsigma_2 \ldots \varsigma_{2\ell-1} \varsigma_{2\ell}. \tag{4.4}$$

Using Lemma 4.3, it can be easily shown that

$$| M(\xi) | = \Pi_{\ell=0}^{n} \int_{-1}^{1} W_\ell^2(x) d\xi(x) = 2^{n(n+1)} \Pi_{\ell=1}^{n} (\varsigma_{2\ell-1}\varsigma_{2\ell})^{n-\ell+1} \tag{4.5}$$

for $q = 1$.

There is a considerable amount of literature concerning the relationship between the sequence of canonical moments $\{p_i\}$ and the corresponding design ξ. (See Studden (1982a, 1982b), Lau (1983)). We state here only those results that are pertinent to some of the D_s-optimal product design problem. The next lemma follows from similar arguments to Lemma 2.3 in Studden (1982a).

Lemma 4.4 *(a) The design corresponding to $(1/2, p_2, 1/2, 1)$ concentrates mass α, $1 - 2\alpha$, α on the points -1, 0, 1, respectively, where $\alpha = p_2/2$. (b) The design corresponding to $(1/2, p_2, 1/2, p_4, 1/2, 1)$ concentrates mass α, $1/2 - \alpha$, $1/2 - \alpha$, α on the points -1, $\sqrt{t}$, $\sqrt{t}$, 1, respectively, where $\alpha = p_2 p_4 / (2(q_2 + p_2 p_4))$, $t = p_2 q_4$. (c) The design corresponding to $(1/2, p_2, 1/2, p_4, 1/2, 1)$ concentrates mass α_1, α_2, $1 - 2\alpha_1 - 2\alpha_2$, α_2, α_1 on the points -1, $\sqrt{t}$, 0, $\sqrt{t}$, 1, respectively, where $\alpha_1 = p_2 p_4 p_6 / (2(1 - t))$, $\alpha_2 = p_2 q_2 q_4 / (2t(1 - t))$, $t = p_2 q_4 + p_4 q_6$.*

5. D_s-optimal product designs.

Let

$$\Xi_1 = \{\xi_1 \times \ \xi_2 \times \ \ldots \times \xi_q : \xi_i \text{ is a design on } [-1, 1]\}.$$

Consider an arbitrary design $\eta = \xi_1 \times \ldots \times \xi_q$ in Ξ_1. For each $j = 1, \ldots, q$, let $W_{i(j)}(x)$ be the orthogonal polynomial of degree i with the leading coefficient 1 with respect to ξ_j. Define

$$N_{\ell,m} = \binom{\ell + m}{m} \tag{5.1}$$

Also denote $M_n(\eta)$ by the information matrix of a design η for the n-th degree polynomial regression model on the q-cube.

Lemma 5.1

$$| M_n(\eta) | = 2^{K_{n,q}} \Pi_{j=1}^{q} \Pi_{i=1}^{n} [\int_{-1}^{1} W_{i(j)}^2(x) d\xi_j(x)]^{N_{q-1,n-i}} \tag{5.2}$$

where $K_{n,q} = 2q \cdot \Sigma_{i=1}^{n} i \cdot N_{q-1,n-i}$.

Proof. Recall that $f(x)$ is the vector of $N_{n,q}$ monomials $x_1^{\ell_1} \ldots x_q^{\ell_q}, \sum_{j=1}^{q} \ell_j \leq n$. Let $g(x)$ be the vector of length $N_{n,q}$ monomials $W_{\ell_1(1)}(x_1) \ldots W_{\ell_1(q)}(x_q)$, $\sum_{j=1}^{q} \ell_j \leq n$. Then it can be easily checked that there exists an $N_{q,n} \times N_{q,n}$ lower triangular matrix A with $| A |= 1$ such that $g(x) = Af(x)$. So

$$\begin{aligned} | M_n(\eta) | &= | \int f(x)\ f(x)'\ d\eta(x) | \\ &= | \int g(x)\ g(x)'\ d\eta(x) | . \end{aligned} \tag{5.3}$$

Note that $\int g(x)\ \ g(x)' d\eta(x)$ is a diagonal matrix since $\int_{-1}^{1} W_{i(j)}(x_j)\ \ W_{\ell(j)}(x_j)$ $\xi_j(dx_j) = 0$ for any j and $\ell \neq i$. Also there exist $N_{q-1,n-i}$ components of $g(x)$ like $W_{i(j_0)}(x_{j_0}) \Pi_{j \neq j_0} W_{\ell_j(j)}(x_j)$ since $\Pi_{j \neq j_0} W_{\ell_j(j)}(x_j)$ is a monomial of degree $\leq n - i$ with $q - 1$ variables. Thus

$$\begin{aligned} | M_n(\eta) | &= \Pi_{\sum_{i=1}^{q} \ell_i \leq n} \Pi_{j=1}^{q} 2^{2\ell_j} \int_{-1}^{1} W_{\ell_j(j)}^2(x_j) d\xi_j(x_j) \\ &= \Pi_{j=1}^{q} \Pi_{i=1}^{n} [2^{2i} \int_{-1}^{1} W_{i(j)}^2(x_j) d\xi_j(x_j)]^{N_{q-1,n-i}} \\ &= 2^{2q \cdot \Sigma_1^q i N_{q-1,n-i}} \cdot \Pi_{j=1}^{q} \Pi_{i=1}^{n} [\int_{-1}^{1} W_{i(j)}^2(x) d\xi_j(x)]^{N_{q-1,n-i}}. \end{aligned}$$

Theorem 5.1 *The D-optimal product design over the class of produce designs Ξ_1 is*

$$\eta_{n,q}^* = \xi_{n,q}^* \times \ldots \times \xi_{n,q}^*, \tag{5.4}$$

in which the canonical moments of $\xi_{n,q}^*$ are given by

$$\begin{aligned} p_{2i-1} &= \frac{1}{2}, \quad i = 1, \ldots, n, \\ p_{2i} &= \frac{q + n - i}{q + 2(n - i)}, \quad i = 1, \ldots, n - 1 \\ \text{and } p_{2n} &= 1. \end{aligned} \tag{5.5}$$

Proof. By Lemma 5.1,

$$| M_n(\eta) |= 2^{K_{n \cdot q}} \Pi_{j=1}^{q} \Pi_{i=1}^{n} \left[\int_{-1}^{1} W_{i(j)}^2(x) d\xi_j(x) \right]^{N_{q-1,n-i}}.$$

Note that

$$\max_{\xi_1,\dots\xi_q} \Pi_{j=1}^{q} \Pi_{i=1}^{n} \left[\int_{-1}^{1} W_{i(j)}^2(x) d\xi_j(x) \right]^{N_{q-1,n-i}}$$
$$= \left[\max_{\xi} \Pi_{i=1}^{n} \left[\int_{-1}^{1} W_i^2(x) d\xi(x) \right]^{N_{q-1,n-1}} \right]^q.$$

So it suffices to find a design $\xi_{n,q}^*$ which maximizes $\Pi_1^n [\int W_i^2(x) d\xi(x)]^{N_{q-1,n-i}}$ and then, the D-optimal product design is

$$\eta_{n,q}^* = \xi_{n,q}^* \times \dots \times \xi_{n,q}^*.$$

It can be easily checked that $N_{q,i} = \sum_{j=0}^{i} N_{q-1,j}$ (Scheffe (1958)). Using this and (4.4), we get

$$\begin{aligned}
&\Pi_{i=1}^{n} [\int_{-1}^{1} W_i^2(x) d\xi(x)]^{N_{q-1,n-i}} \\
&= (\varsigma_1\varsigma_2)^{N_{q-1,n-1}} (\varsigma_1\varsigma_2\varsigma_3\varsigma_4)^{N_{q-1,n-2}} \dots (\varsigma_1\varsigma_2 \dots \varsigma_{2n-1}\varsigma_{2n})^{N_{q-1,0}} \\
&= (\varsigma_1\varsigma_2)^{N_{q,n-1}} (\varsigma_3\varsigma_4)^{N_{q,n-2}} \dots (\varsigma_{2n-1}\varsigma_{2n})^{N_{q,0}} \\
&= (p_1 q_1 p_2)^{N_{q,n-1}} (q_2 p_3 q_3 p_4)^{N_{q,n-2}} \dots (q_{2n-2} p_{2n-1} q_{2n-1} p_{2n})^{N_{q,0}} \qquad (5.6)
\end{aligned}$$

Simple algebra shows that (5.6) is maximized at

$$\begin{aligned}
p_{2i-1} &= \frac{1}{2}, \quad i = 1, \dots n, \\
p_{2i} &= \frac{N_{q,n-i}}{N_{q,n-i} + N_{q,n-(i+1)}} = \frac{q+n-i}{q+2(n-i)}, \quad i = 1, \dots, n-1, \\
p_{2n} &= 1.
\end{aligned}$$

The uniqueness of the D-optimal product design comes from $p_{2n} = 1$.

For $q = 2$ case, we get $p_{2i} = (n-i+2)/2(n-i+1)$, $p_{2i-1} = 1/2$ and $p_{2n} = 1$ from Theorem 5.1. In the following examples we illustrate the D-optimal product design for $2 \le n \le 4$ and $q = 2$ by using Lemma 4.4.

Example 5.1 Suppose $n = 2$. Then $p_2 = 3/4$ and $p_4 = 1$. By Lemma 4.4, the corresponding design is

$$\xi_{2,2}^* = \begin{bmatrix} -1 & 0 & 1 \\ \frac{3}{8} & \frac{1}{4} & \frac{3}{8} \end{bmatrix}$$

and $\eta^*_{2,2} = \xi^*_{2,2} \times \xi^*_{2,2}$ is the D-optimal product design.

Example 5.2. Suppose $n = 3$. Then $p_2 = 2/3$, $p_4 = 3/4$ and $p_6 = 1$. By Lemma 4.4,

$$\xi^*_{3,2} = \begin{bmatrix} -1 & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & 1 \\ .3 & .2 & .2 & .3 \end{bmatrix}$$

and $\eta^*_{3,2} = \xi^*_{3,2} \times \xi^*_{3,2}$ is the D-optimal product design.

Example 5.3. Suppose $n = 4$. Then $p_2 = 5/8$, $p_4 = 2/3$, $p_6 = 3/4$ and $p_8 = 1$. By Lemma 4.4,

$$\xi^*_{4,2} = \begin{bmatrix} -1 & -\sqrt{\frac{9}{24}} & 0 & \sqrt{\frac{9}{26}} & 1 \\ \frac{1}{4} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{4} \end{bmatrix}$$

and $\eta^*_{4,2} = \xi^*_{4,2} \times \xi^*_{4,2}$ is the D-optimal product design.

We consider the usual D-efficiency defined by

$$D(\xi) = \left[\frac{|\,M(\xi)\,|}{|\,M(\xi^*)\,|}\right]^{\frac{1}{K}}, \tag{5.7}$$

where ξ^* is a D-optimal design,and the G-efficiency defined by

$$G(\xi) = \frac{K}{\sup_x d(x, \xi)} \tag{5.8}$$

to see how good D-optimal product designs are for $q = 2$.

Let $R_i(x)$ be the orthonormal polynomial of degree i with respect to $\xi^*_{n,2}$. By using (4.3) and (4.4) it can be easily checked that

$$\begin{aligned} &R_0(x) = 1, \quad R_1(x) = \frac{1}{2\sqrt{\varsigma_1\varsigma_2}} \cdot x \\ \text{and} \quad &R_\ell(x) = x \cdot \frac{1}{2}\frac{1}{\sqrt{\varsigma_{2\ell-1}\varsigma_{2\ell}}} R_{\ell-1}(x) - \sqrt{\frac{\varsigma_{2\ell-3}\varsigma_{2\ell-2}}{\varsigma_{2\ell-1}\varsigma_{2\ell}}} R_{\ell-2}(x), \quad \ell \geq 2. \end{aligned} \tag{5.9}$$

From (4.2) and (5.5), $\varsigma_1\varsigma_2 = (n+1)/2n$ and $\varsigma_{2\ell-1}\varsigma_{2\ell} = 1/16$ for $\ell \geq 2$. So

$$\begin{aligned} &R_0(x) = 1, \quad R_1(x) = \sqrt{\frac{2n}{n+1}}x, \quad R_2(x) = 2 \cdot \sqrt{\frac{2n}{n+1}}x^2 - 2 \cdot \sqrt{\frac{n+1}{2n}} \\ \text{and} \quad &R_\ell(x) = 2xR_{\ell-1}(x) - R_{\ell-2}(x), \quad \ell \geq 3. \end{aligned} \tag{5.10}$$

Since the variance function is invariant under linear nonsingular transformation of $f(x)$, $d(x, \eta^*_{n,2})$ can be written as

$$d(x, \eta^*_{n,2}) = \sum_{\substack{0 \leq i \leq j \leq n \\ i+j \leq n}} R_i^2(x_1) R_j^2(x_2), \tag{5.11}$$

which can be calculated easily by using the recursive relations (5.10). Sup $d(x, \eta^*_{n,2})$ was computed numerically on VAX 11/780. The D- efficiency for $3 \le n \le 5$ was based on a numerical D-optimal design which was found in the previous section. As mentioned in Section 3, the optimization algorithm would not produce numerical D-optimal designs for $6 \le n \le 12$. Therefore we could not get values for the D-efficiency in these cases. However, by using the inequality

$$D(\xi) \ge \exp\left\{-\frac{1-G(\xi)}{G(\xi)}\right\} \tag{5.12}$$

in Kiefer (1962b), a lower bound of the D-efficiency for $6 \le n \le 12$ is given in Table 5.1.

Table 5.1 Efficiency of D-optimal product designs when $q = 2$.

degree n	$(x^0_{1,n}, x^0_{2,n})^*$	$d(x^0_{1,n}, x^0_{2,n})$	K	G-efficiency	D-efficiency
2	(0,0)	7.000	6	.8571	.9952
3	(1,.3103)	10.2260	10	.9779	.9937
4	(0,0)	17.2500	15	.8696	.9922
5	(1,.6989)	22.1270	21	.9491	.9928
6	(0,0)	31.3333	28	.8936	.8878**
7	(1,.8366)	38.0338	36	.9465	.9451**
8	(0,0)	49.3750	45	.9114	.9074**
9	(1,.8980)	58.0581	55	.9473	.9458**
10	(0,0)	71.4000	66	.9244	.9214**
11	(1.,9303)	81.2191	78	.9487	.9596**
12	(0,0)	97.4167	91	.9341	.9319**

$^*d(x^0_{1,n}, x^0_{2,n}) = \sup_{x_1,x_2} d((x_1, x_2), \eta^*_{n,2})$
**a lower bound

In the case where interest is in only the $(n-m)$ highest order terms, i.e., $(m+1)$-th, $\ldots, n$-th degree terms, we give a similar analysis.

Theorem 5.2 The D_s-optimal product design for the $(n-m)$ highest order terms over the class of product designs Ξ_1 is

$$\eta^*_s = \xi^*_s \times \ldots \times \xi^*_s,$$

in which the canonical moments of ξ^*_s are

$$\begin{aligned} p_{2i-1} &= \frac{1}{2}, \quad i = 1, \ldots, n, \\ p_{2i} &= \frac{N_{q,n-i} - N_{q,m-i}}{N_{q,n-i} - N_{q,m-i} + N_{q,n-i-1} - N_{q,m-i-1}}, \quad i = 1, \ldots, m, \\ p_{2i} &= \frac{q+n-i}{q+2(n-i)}, \quad i = m+1, \ldots, n-1, \end{aligned} \tag{5.13}$$

and $p_{2n} = 1$. Here $N_{\ell,m}$ is given by (5.1).

Proof Recall the computation formula for $|\Sigma(\eta)|$ and use (5.2). Then

$$\begin{aligned}|\Sigma(\eta)| =& 2^{K_n-K_m}\cdot\frac{|M_n(\eta)|}{|M_m(\eta)|}\\ =& 2^{K_n-K_m}\cdot\frac{\prod_{j=1}^q\prod_{i=1}^n[\int W_{i(j)}^2(x)d\xi_j(x)]^{N_{q-1,n-i}}}{\prod_{j=1}^q\prod_{i=1}^m[\int W_{i(j)}^2(x)d\xi_j(x)]^{N_{q-1,m-1}}}\\ =& 2^{K_n-K_m}\cdot\prod_{j=1}^q\Bigg[[\prod_{i=1}^m\int W_{i(j)}^2(x)d\xi_j(x)]^{N_{q-1,n-i}-N_{q-1,m-i}}\cdot\\ &\prod_{i=m+1}^n[\int W_{i(j)}^2(x)\xi_j(dx)]^{N_{q-1,n-i}}\Bigg] \qquad (5.14)\end{aligned}$$

The rest is the analogous to the D-optimal produce design case.

As special cases, first we consider $m = n-1$, i.e., all the highest order terms. By the substitution of $m = n-1$ into (5.13) and then, simplification of the resulting expression, we get

$$\begin{aligned} p_{2i} =& \frac{q-1+n-i}{q-1+2(n-i)}, \quad 1\le i\le n-1 \qquad (5.15)\\ \text{and} \quad p_{2n} =& 1.\end{aligned}$$

For the $q = 2$ case, interestingly the canonical moments correspond to the D-optimal design for the n-th degree polynomial regression on [-1,1]. For $m = n-2$, i.e., all the highest and second highest terms, (5.13) is simplified to

$$\begin{aligned} p_{2i} =& \frac{2(n-i)^2+3(q-5)(n-i)+(q-1)(q-2)}{4(n-i)^2+4(q-2)(n-i)+(q-1)(q-2),} \quad 1\le i\le n-2,\\ p_{2(n-1)} =& \frac{q+1}{q+2} \quad \text{and } p_{2n}=1. \qquad (5.16)\end{aligned}$$

For $q = 2$ case; (5.16) reduces to $p_{2i} = \big(2(n-i)+1\big)/4(n-i)$, $1\le i\le n-1$ and $p_{2n} = 1$.

The D_s-efficiency of the product design for $q = 2$, $m = n-1$, $3\le n\le 5$ are .9727, .9569, .9605, respectively. Also the D_s-efficiency for $q = 2$, $m = 1$, $n = 3$ is .9902. All the D_s-efficiency are based on numerical D_s-optimal designs in section 3.

References

[1] C. L. Atwood. (1969). Optimal and efficient designs of Experiments. *Ann. Math. Statist.*, Vol. 40, 1570-1602.

[2] R. H. Farrell, J. Kiefer and J. Walbran. (1967). Optimum multivariate designs. *Proc. Fifth Berkeley Symp.*, Vol. I, Univ. of California Press, 113-138.

[3] V. V. Federov. (1972). *Theory of optimal experiments.* Translated by W. J. Studden and E. M. Klimko, Academic Press, New York.

[4] A. Giovagnoli, F. Pukelsheim and H. Wynn. (1987). Group invariant orderings and experimental designs. (To appear in *J. Statist. Plann. Inference*).

[5] S. Karlin and W. J. Studden. (1966a). Tchebycheff systems: with applications in analysis and statistics. *Interscience Publisher*, New York.

[6] S. Karlin and W. J. Studden. (1966b). Optimal experimental designs. *Ann. Math. Statist.*, Vol. 37, 783-815.

[7] J. Kiefer. (1959). Optimum Experimental Designs. *J. Roy. Statist. Soc. Ser. B*, Vol. 32, 272-319.

[8] J. Kiefer. (1961a). Optimum designs in regression problems, II. *Ann. Math. Statist.*, Vol. 32, 298-325.

[9] J. Kiefer. (1961b). Optimum experimental design V, with applications to systematic and rotatable designs. *Proc. Fourth Berkeley Symp.*, Vol. I, Univ. of California Press, 381-405.

[10] J. Kiefer. (1974). General equivalence theory for optimum designs (approximation theory). *Ann. Statist.*, Vol. 2, 849-879.

[11] J. Kiefer and J. Wolfowitz. (1959). Optimum designs in regression problem. *Ann. Math. Statist.*, Vol. 30, 271-294.

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

[13] K. Kono. (1962). Optimal designs for quadratic regression on the k-cube. *Mem. Fac. Sci. Kynshu Univ. Ser. A.*, Vol. 16, 114-122.

[14] T. S. Lau. (1983). Theory of canonical moments and its application in polynomial regression. Technical Report 83-23, Purdue University.

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

[16] Y. B. Lim. (1986). Optimal designs in multivariate polynomial regression and polynomial spline regression. Ph.D. Thesis, Purdue University.

[17] Y. B. Lim, W. J. Studden and H. P. Wynn. (1986). A Note on approximate D-optimal designs for $G \times 2^m$. (to appear in *Statistical Decision Theory and Related Topics IV*).

[18] H. Scheffe. (1958). Experiments with mixtures. *J. Roy. Statist. Soc. B*, Vol. 20, 344-360.

[19] W. J. Studden. (1980). D_S-optimal designs for polynomial regression using continued fractions. *Ann. Statist.*, Vol. 8, 1132-1141.

[20] W. J. Studden. (1982a). Some robust-type D-optimal design in polynomial regression. *J. Amer. Statist. Assoc.*, Vol. 77, 916-921.

[21] W. J. Studden. (1982b). Optimal designs for weighted polynomial regression using canonical moments. *Statistical Decision Theory and Related Topics III*, Academic Press, New York.

[22] W. Welch. (1985). ACED: Algorithms for the Construction of Experimental Designs. *The American Statistician*, Vol. 39, 146.