OPTIMUM MULTIVARIATE DESIGNS

R. H. FARRELL,[1] J. KIEFER,[1] and A. WALBRAN
CORNELL UNIVERSITY

1. Introduction

1.1. *Notation and preliminaries.* This paper is concerned with the computation of optimum designs in certain multivariate polynomial regression settings.

Let $f = (f_1, \cdots, f_k)$ be a vector of k real-valued continuous linearly independent functions on a compact set X. We shall work in the realm of the approximate theory discussed in many of the references, wherein a design is a probability measure ξ (which can be taken to be discrete) on X. The information matrix $M(\xi)$ of the design ξ for problems where the regression function is $\sum_1^k \theta_i f_i(x)$ (with $\theta = (\theta_1, \cdots, \theta_k)$ unknown and with uncorrelated homoscedastic observations and quadratic loss considerations of best linear unbiased estimators) has elements $m_{ij}(\xi) = \int f_i f_j \, d\xi$. Thus, det $M^{-1}(\xi)$ is proportional to the generalized variance of the best linear estimators of all θ_i. We denote by Γ the space of all $M(\xi)$. We shall have occasion to consider the set of all *distinct* functions of the form $f_i f_j$, $i \geq j$, and shall write them as $\{\phi_t, 1 \leq t \leq p\}$. We then write $\mu_t(\xi) = \int \phi_t \, d\xi$. Whether or not some ϕ_t is a nonzero constant (as it is in our polynomial examples), we define $\phi_0(x) \equiv 1$ and $\mu_0 = 1$.

The main results of this paper characterize, for certain X and f, some designs ξ^* which are D-optimum; that is, for which

$$\det M(\xi^*) = \max_{\xi} \det M(\xi). \tag{1.1}$$

Define, for $M(\xi)$ nonsingular,

$$\begin{aligned} d(x, \xi) &= f(x)M^{-1}(\xi)f(x)', \\ \bar{d}(\xi) &= \max_{x \in X} d(x, \xi). \end{aligned} \tag{1.2}$$

The quantity $d(x, \xi)$ is proportional to the variance of the best linear estimator of the regression $f(x)\theta'$ at x. A result of Kiefer and Wolfowitz [8] is that ξ^* satisfies (1.1) if and only if it satisfies the G-(global-) optimality criterion

$$\bar{d}(\xi^*) = \min_{\xi} \bar{d}(\xi), \tag{1.3}$$

and that (1.1) and (1.3) are satisfied if and only if

$$\bar{d}(\xi^*) = k. \tag{1.4}$$

If the support of an optimum design is exactly k points, then ξ is uniform on those points. Our main way of finding D- and G-optimum (hereafter simply

[1] Research supported by ONR Contract Nonr-401(03).

called "optimum") designs and of verifying their optimality is thus to guess a ξ^* (perhaps by minimizing det $M(\xi)$ over some subset of designs depending on only a few parameters) and then to verify (1.4). We also record here the fact that all optimum ξ^* have the same $M(\xi^*)$, and that they all satisfy

$$\xi^*(\{x: d(x, \xi^*) = k\}) = 1. \tag{1.5}$$

It is often the case that there is a compact group $G = \{g\}$ of transformations on X, with associated transformations $\{\bar{g}\}$ on $\{\xi\}$, and such that $d(gx, \xi) = d(x, \bar{g}\xi)$. In such a case (see Kiefer [6]), there is G-invariant optimum design ξ^* (that is, such that $\xi^*(gA) = \xi^*(A)$ for all g and A), and the function $d(\cdot, \xi^*)$ and set of (1.5) are G-invariant.

Whereas some of our discussion refers to general X and f, our detailed examples of sections 2, 3, and 4 all treat problems of *polynomial regression in q variables, of degree $\leq m$.* Here X is a compact q-dimensional Euclidean set whose points we usually denote by $x = (x_1, \cdots, x_q)$, and the $f_i(x)$ are of the form $\prod_{j=1}^{q} x_j^{m_j}$ where the m_j are nonnegative integers with sum $\leq m$. It is well known in this case that

$$k = \binom{m+q}{q}. \tag{1.6}$$

Moreover, since the $f_i f_j$ are all the monomials of degree $\leq 2m$, we have

$$p = \binom{2m+q}{q}. \tag{1.7}$$

The three examples we shall treat in detail are (in section 4) the unit q-ball

$$\left\{x: \sum_1^q x_i^2 \leq 1\right\}; \tag{1.8}$$

(in section 3) the q-cube

$$\{x: \max_{1 \leq i \leq q} |x_i| \leq 1\}; \tag{1.9}$$

and (in section 2) the unit simplex, which it is more convenient to represent in barycentric coordinates $x = (x_0, x_1, \cdots, x_q)$ as

$$\left\{x: \min_{0 \leq i \leq q} x_i \geq 0, \sum_0^q x_i = 1\right\}. \tag{1.10}$$

These are perhaps the three generalizations which are simplest, most natural, and of greatest practical importance, of the unit interval ($q = 1$), which is discussed in section 2. Unfortunately, the simple structure which is present when $q = 1$ and which is reflected in the elegant results of Guest [3] and Hoel [4] does not carry over to $q > 1$, and the results depend strongly on the shape of X; even in the case of the simplex where at least some analogous results seem to hold, they cannot be obtained by the same methods. We now indicate how this is reflected in the geometry of Γ.

1.2. *The geometry of* Γ. The set Γ can clearly be regarded as a convex body in p-dimensional Euclidean space with coordinates μ_t, $1 \leq t \leq p$; of course,

$p \leq k(k+1)/2$. Write $a = (a_0, a_1, \cdots, a_p)$. Let $\sum_1^p a_t\mu_t + a_0 = 0$ be a supporting hyperplane of Γ with $\sum_1^p a_t\mu_t + a_0 \geq 0$ in Γ. Clearly, the supporting polynomial $T(x; a) = \sum_1^p a_t\phi_t(x) + a_0$ is nonnegative on X.

(For future reference, the reader should note in connection with the previous and next paragraphs that, if ξ^* is optimum and $\gamma^* = M(\xi^*)$, then ξ^* is admissible and hence γ^* is a boundary point, and $k - d(x, \xi^*)$ supports Γ at γ^*.)

Let $\gamma_0 = M(\xi_0)$ be a boundary point of Γ. A supporting polynomial $T(\cdot\,; a^0)$ which supports Γ at γ_0 is then ≥ 0 on X and is 0 on the support of ξ_0. Thus, an analysis of what the set of zeros of a T of the above form can be can yield information about the boundary points of Γ (while the extreme points are clearly a subset of the points corresponding to ξ's with one-point support). For example, in the well-known univariate polynomial case $X = [0, 1]$, $k = m + 1$, $f_i(x) = x^{i-1}$, any such T is a nonnegative polynomial on X of degree $\leq 2m$, which (if not identically zero) therefore has at most $m + 1$ zeros, at most m of which are in the interior of X. In this example, moreover, if $\gamma = M(\xi)$ is an arbitrary point of Γ and $\xi^{(0)}(0) = 1$, the line from the boundary point $M(\xi^{(0)})$ through γ passes through another boundary point $\gamma' = M(\xi')$, so that $\gamma = M(\lambda\xi^{(0)} + (1 - \lambda)\xi')$ with $0 \leq \lambda \leq 1$; thus one concludes that any point of Γ can be represented as $M(\xi'')$ for a ξ'' supported by at most $m + 1$ points. One can also characterize the admissible ξ easily in this example as the boundary points with at most $m - 1$ points of support in the interior of X (Kiefer [5]).

Unfortunately, the examples studied in the present paper (as well as non-polynomial, and especially non-Chebyshev systems in one dimension) do not yield such simple analyses. This is clear when one considers the more complex sets on which a T can now vanish. For example, in the case of linear regression ($m = 1$) on the square (1.9) with $q = 2$, any supporting T which is not identically zero, being quadratic, vanishes either on a subset of the corners of X, or at a single point of X, or on a line of X. In the latter case we invoke the one-dimensional result to conclude that at most two points are needed to support a ξ yielding this $M(\xi)$; thus, every boundary point of Γ is obtainable from a ξ supported either by a subset of the corners or else by at most two other points. Replacing $\xi^{(0)}$ in the argument of the previous paragraph by the measure which assigns all probability to the point $(-1, -1)$, we conclude that every point of Γ can be obtained from a ξ which is supported either by a subset of the corners or else by at most three points of X. If we replace (1.9) by (1.10) with $q = 2$, we obtain that at most 3 points rather than 4 are needed. The admissible points can be characterized similarly, but it is clear that the difficulty of obtaining such characterizations will be much greater for larger q and m. As for the optimum design, it is the uniform distribution on the 3 corners in the case (1.10), on the 4 corners in the case (1.9), and, for another example, on the 5 corners if X is a symmetric pentagon. The uniqueness in all three cases can be proved by the method given in the next paragraph, but in other cases, such as (1.8), there is no uniqueness. Section 3.3 of [6] characterizes optimum designs for linear regression on general compact X in q dimensions.

The increased complexity in higher dimensions is also present in the uniqueness question: given $\gamma = M(\xi^*)$, when is there no other ξ with $M(\xi) = \gamma$? This can sometimes be answered as follows. Suppose γ is a boundary point and that a supporting polynomial T at γ has exactly L zeros $x^{(1)}, x^{(2)}, \cdots, x^{(L)}$ on X. Any ξ with $M(\xi) = \gamma$ must be supported by a subset of $\{x^{(1)}, \cdots, x^{(L)}\}$, and must satisfy $\sum_j \phi_t(x^{(j)})[\xi^*(x^{(j)}) - \xi(x^{(j)})] = 0$ for $0 \leq t \leq p$. Hence, if rank $\{\phi_t(x^{(j)}), 0 \leq t \leq p, 1 \leq j \leq L\} = L$, then ξ^* is the unique design yielding γ. In the univariate polynomial example of the second paragraph above, each boundary point γ can be proved by this device to be yielded by a unique ξ^*.

The prescription outlined just below (1.4) for verifying optimality, and which has worked well when $q = 1$ or $m \leq 2$, is difficult to apply in other cases. This is because $k - d(x, \xi^*)$ can no longer be written as a sum of a small number of obviously nonnegative simple polynomials, but may instead require a large number of rational functions for such a representation. The decision procedures (Tarski, Henkin, and others) for representing or verifying nonnegativity of such polynomials are unwieldy to implement in these problems. The example of the simplex (1.10) with $q = 2$, $m = 3$, treated in section 2 by direct analysis, illustrates the increased complexity. In other cases we have been unable to obtain analytical verifications of optimality and have used machine methods to obtain results which are satisfactory from a practical point of view but which, theoretically, only yield statements of results which hold to within a certain accuracy, rather than complete proofs of the exact results.

We end this subsection with a simple observation which is often useful in optimum design theory for polynomial regression on a q-dimensional set X. If B is a subset of X such that for some $q \times q$ orthogonal matrix A and some scalar b with $|b| > 1$, the set $bAB = \{x: b^{-1}A^{-1}x \in B\}$ is also a subset of X, then no design ξ supported by B can be optimum. This follows at once upon defining ξ' by $\xi'(C) = \xi(b^{-1}A^{-1}C)$ for $C \subset X$ and computing $\det M(\xi') = b^k \det M(\xi)$. In particular, if X is such that $x \in X \Rightarrow ax \in X$ for $0 \leq a < 1$, then the support of any optimum design must contain at least one point of the boundary of X. The considerations of this paragraph can be modified in an obvious way for admissibility questions.

1.3. *Number of points needed for an optimum design.* An aspect of the geometry of Γ which is of particular practical importance is the minimum number N of points such that there is an optimum design supported by N points. (It will be clear how to modify much of the discussion which follows to treat this question for points of Γ other than those corresponding to D-optimum designs, but for brevity we will treat only the latter.) An optimum design will be called *minimal* if no proper subset of its support is the support of an optimum design. We shall see that this property is broader than that of being an optimum design on N points; the latter will be called *absolutely minimal.*

Clearly $N \geq k$. On the other hand, if there is a matrix B of rank b such that $\sum_{j=1}^{p} b_{ij}\phi_j(x)$ is a constant function of x for each i, then Γ has dimension $\leq p - b$.

Since the extreme points of Γ can be obtained from ξ's with one-point support, we obtain the trivial bounds

$$k \leq N \leq \min(p - b + 1, k(k + 1)/2), \tag{1.11}$$

where the well-known bound $k(k + 1)/2$, which is relevant only when $p = k(k + 1)/2$ and $b = 0$, is a consequence of the fact that optimum designs correspond to certain boundary points of Γ. In the polynomial case we have $b = 1$ (since 1 is a ϕ_t) and thus, from (1.6) and (1.7),

$$\binom{m + q}{q} \leq N \leq \binom{2m + q}{q}. \tag{1.12}$$

Of greater use is the upper bound one can obtain once one knows some optimum design ξ^*. Let $V = \{x\colon d(x, \xi^*) = k\}$ (see (1.5)), and let W be the support of ξ^*. We denote the number of points in these sets by v and w. (When v or w is infinite, as in the example of the ball (1.8) for $q \geq 2$ in section 4, it is easy to see that (1.14) below still holds, but we shall usually treat the finite case.) Let $U = V$ or W (and $u = v$ or w). The $p + 1$ linear equations

$$\sum_{x \in U} \phi_t(x)\xi(x) = \mu_t(\xi^*), \qquad 0 \leq t \leq p \tag{1.13}$$

in the unknowns $\xi(x)$ are consistent (since $\{\xi^*(x)\}$ is a solution), so that the dimensionality of the linear set H (say) of solutions of (1.13) is $u - h$ where $h = \text{rank}\ \{\phi_t(x), 0 \leq t \leq p, x \in U\}$.

Considering H as a set in the u-dimensional space with coordinates $\xi(x)$, $x \in U$, we know that ξ^*, with all coordinates nonnegative, is in H, and conclude easily that H contains a point with all coordinates nonnegative and with at least h zero coordinates. Hence,

$$N \leq \text{rank}\ \{\phi_t(x), 0 \leq t \leq p, x \in U\}. \tag{1.14}$$

(Of course, if U is replaced by X, this becomes the $p - b + 1$ of (1.11).) We will illustrate the use of this in the polynomial case (where $1 \in \{\phi_t\}$ so that the domain of t in (1.14) can be taken to be $1 \leq t \leq p$) in section 3.3, in the case of quadratic regression on the q-cube (1.9). We can think of such polynomial applications in terms of finding a matrix C of rank c such that $\sum_{j=1}^{p} c_{ij}\phi_j(x) \equiv 0$ on W. We can then conclude that, if N_W is the minimum number of points in W supporting an optimum design, then (paralleling (1.11))

$$N_W \leq p - c. \tag{1.15}$$

On the other hand, it is obvious from (1.13) that, for $U = V$ or W,

$$N_U = u \quad \text{if rank} \quad \{\phi_t(x), 0 \leq t \leq p, x \in U\} = u. \tag{1.16}$$

It is easy to give examples which illustrate the fact that we can have $N_W > N$; that is, that minimality and absolute minimality do not coincide. For example, in the case $m = 1$, $q = 2$ of linear regression on the unit disc (1.8), the discussion of the fourth paragraph of section 1.2 shows that the set of j equally spaced points on the boundary is minimal if $j = 3$, 4, or 5, but is absolutely minimal only when $j = 3$.

Also, the designs whose optimality is easiest to verify are often ones which are symmetric, that is, invariant in the sense described below (1.5), and there is no reason why minimal designs should be of this form. For example, in the case of the q-cube (1.9) with $m = 1$, the uniform distribution on the 2^q corners is the only optimum design invariant under the symmetries of the q-cube, but if q is such that there exists a $(q + 1) \times (q + 1)$ Hadamard matrix (for example, if $q + 1$ is a power of 2), then it is well known that there is an optimum design on $k = q + 1$ corners (namely, the corners of an inscribed regular q-simplex). This example illustrates another technique for reducing an upper bound on N or N_W; in sections 3 and 4 we shall see how the use of various known results on orthogonal arrays and rotatable configurations can be used similarly.

The search for absolutely minimal designs can be described as a programming problem, of finding a nonnegative solution of (1.13) with $U = V$, which has a minimum number of nonzero elements. Analytical or machine methods for solving this problem would seem important.

2. The simplex

We have mentioned in section 1 that this case (1.10) evidences the most regular mathematical behavior among q-dimensional sets X. In the linear and quadratic cases it has been known for some time that the simplex exhibits a behavior (described precisely, below) very much like that present when $q = 1$. This phenomenon appears to carry over to cubic and perhaps higher degree regression, although we have as yet proved only one small fragment of the conjectured general result, and have machine computations in only two other cases. To describe these results, let E_m be the set of $m + 1$ points supporting the Guest-Hoel design when $q = 1$. (Thus, $E_1 = \{x_0 = 0, 1\}$; $E_2 = \{x_0 = 0, \frac{1}{2}, 1\}$; $E_3 = \{x_0 = 0, (1 \pm 5^{-1/2})/2, 1\}$; $E_4 = \{x_0 = 0, \frac{1}{2}, 1, (1 \pm (\frac{3}{7})^{1/2})/2\}$; and so on.)

The results in the linear and quadratic cases for general dimension q (Kiefer [7]) can be summarized by stating that, for degrees $m = 1$ and 2, the unique optimum design assigns equal measure to each of the points which is in the E_m of some edge of the q-simplex (when that edge is considered as a 1-simplex). We cannot hope for this pattern for $m > 2$, since the E_m points on all edges will be fewer in number than the k of (1.6). However, one can still conjecture that one or all of the following are true: (1) there is an optimum design whose support includes the E_m points on all edges (and no other points on edges); (2) there is an optimum design which assigns equal measure to each of k points; (3) the optimum design is unique; (4) generalizing the vertex- and edge-stationarity of (1), for fixed m there are optimum designs for dimension q which have the same support on the r-dimensional faces of X for $q \geq r$; (5) the design of (4) has points of support only on faces of dimension $\leq \min(m - 1, q)$.

What we have succeeded in treating analytically is the case $m = 3$, $q = 2$, and the details will be found at the end of this section. We have also observed, by machine search, that (a) the optimum design in the case $m = 3$, $q = 3$ appears to give equal weights to the E_3-points on edges (including vertices) and the midpoints of 2-dimensional faces, just as in the cases $m = 3$, $q = 1$ or 2; (b) the optimum design in the case $m = 4$, $q = 2$ appears to give equal weights to the E_4-points on edges and to the three points of the form $\{x_h = 0.567, x_i = x_j\}$. One can also prove analytically for $m = 3$ and general k that, among all designs which assign equal weights to the vertices, midpoints of 2-dimensional faces, and points on edges satisfying $x_i = b = 1 - x_j$, the choice $b = (1 \pm 5^{-1/2})/2$ minimizes the generalized variance for each k. (The generalized variance for such designs for $q \geq 2$ is proportional to $[v(1 - 2b)^{1/2}]^{-2q(q+1)}$, where $v = b(1 - b)$.) All of the above results conform with the conjectures of the previous paragraph.

If any of the general conjectures are true, they would constitute a deep new result in the area of multidimensional moment and approximation theory. Evidently a new approach is needed, perhaps even to verify analytically (b) and (c) of the previous paragraph. The technique employed for low dimensions and/or degrees, for example, by Kiefer [7] and Uranisi [11], has been that described at the end of section 1.2, and the difficulties encountered for larger q or m are as described there. Even in the case $m = 3$, $q = 2$ which we now consider, a much more brutal approach is used, and it does not suffice when $q = 3$.

THEOREM 2.1. *For $m = 3$, $q = 2$, the unique optimum design ξ^* assigns measure $\frac{1}{10}$ to each of the three vertices, the point $x_0 = x_1 = x_2 = \frac{1}{3}$, and the six points $\{x_h = 0, x_i = 1 - x_j = (1 + 5^{-1/2})/2.\}$.*

PROOF. We shall show that $10 - d(x, \xi^*) \geq 0$ on X, with equality only at the ten points supporting ξ^*. Since the function d is the same for all optimum designs, any optimum design must have this same support, and the weights are unique since there are 10 points and 10 functions. This yields uniqueness.

It is convenient to consider, in place of the coordinates x_i, the coordinates $\beta, t(-\frac{1}{2} \leq \beta \leq 1, 0 \leq t \leq 1)$ satisfying $3x_0 = 1 - t, 3x_1 = 1 - \beta t, 3x_2 = 1 + (\beta + 1)t$ on the portion $0 \leq x_0 \leq x_1 \leq x_2$ of X which, because of the symmetry of ξ^*, is all we need consider. For fixed β, variation of t from 0 to 1 yields a segment from center to edge of X. Write $L = \beta^2 + \beta + 1$ (so that $\frac{3}{4} \leq L \leq 3$). A simple computation yields

$$
\begin{aligned}
9 \sum_{i<j} x_i x_j &= 3 - Lt^2, \\
81 \sum_{i<j} x_i^2 x_j^2 &= 3 - 6(L - 1)t^3 + L^2 t^4, \\
729 \sum_{i<j} x_i^3 x_j^3 &= 3 + 3Lt^2 - 21(L - 1)t^3 + 3L^2 t^4 \\
&\quad + 3L(L - 1)t^5 + (-L^3 + 3L^2 - 6L + 3)t^6, \\
27 \prod_i x_i &= 1 - Lt^2 + (L - 1)t^3.
\end{aligned}
\tag{2.1}
$$

A straightforward computation of $M(\xi^*)$ and $d(x, \xi^*)$ (for example, in terms of the functions x_i, $x_i x_j$, $x_i x_j(x_i - x_j)$, and $\Pi_i x_i$, with $i < j$) yields

$$1 - d(x, \xi^*)/10 = 12 \sum_{i<j} x_i x_j - 120 \sum_{i<j} x_i^2 x_j^2 + 300 \sum_{i<j} x_i^3 x_j^3 + \prod_i x_i[-102 + 410 \sum_{i<j} x_i x_j - 1512 \prod_i x_i]. \tag{2.2}$$

From (2.1) and (2.2) we have, writing $g(t, L) = 729[1 - d(x, \xi^*)/10]/6t^2$,

$$g(t, L) = 131L - 318(L - 1)t - 77L^2t^2 + 449L(L - 1)t^3 - [50L^3 + 102(L - 1)^2]t^4. \tag{2.3}$$

We must show $g(t, L) \geq 0$ for $0 < t \leq 1$, $\frac{3}{4} \leq L \leq 3$. We note that $g(1, L) = 2(3 - L)(5L - 6)^2$, so that the zeros of g on the boundary of X are precisely the vertex ($L = 3$) and E_3-point $L = \frac{6}{5}$.

Writing $f(t, L) = [g(t, L) - g(1, L)]/(1 - t)$, we obtain

$$\begin{aligned} f(t, L) &= [50L^3 - 270L^2 + 563L - 216] \\ &\quad + [50L^3 - 270L^2 + 245L + 102]t \\ &\quad + [50L^3 - 347L^2 + 245L + 102]t^2 + [50L^3 + 102(L - 1)^2]t^3 \\ &= D(L) + C(L)t + B(L)t^2 + A(L)t^3 \quad \text{(say)}. \end{aligned} \tag{2.4}$$

We shall show that $f > 0$ for $0 \leq t \leq 1$, $\frac{3}{4} \leq L \leq 3$, and this will complete the proof.

We have $D(\frac{3}{4}) > 0$ and $D'(L) = 150L^2 - 540L + 563 > 0$. Thus,

$$A(L) > 0, \qquad D(L) > 0, \qquad \tfrac{3}{4} \leq L \leq 3. \tag{2.5}$$

Also, one sees easily that, for $\frac{3}{4} \leq L \leq 1$, we have $B(L) \geq 50L^3$ and $C(L) \geq 0$. We conclude that $f(t, L) > 0$ for $0 \leq t \leq 1$, $\frac{3}{4} \leq L \leq 1$. We divide the region $1 \leq L \leq 3$ into two parts, the division λ being the zero of $C(L)$ in $1 \leq L \leq 3$ ($1.6 < \lambda < 1.7$).

For $\lambda \leq L \leq 3$, we have $C(L) \leq 0$. We shall show the positivity of something $\leq f$, namely,

$$\begin{aligned} f(t, L) + C(L)(1 - t)^2(1 + t) &= [100L^3 - 540L^2 + 808L - 114] \\ &\quad - 77L^2t^2 + [100L^3 - 168L^2 + 41L + 204]t^3 \\ &= E(L) - 77L^2t^2 + F(L)t^3 \text{ (say)}. \end{aligned} \tag{2.6}$$

We first note that $-23L^2 + 132L - 114$ is positive at $L = 1.6$ and $L = 3$ and, hence, for $\lambda \leq L \leq 3$. Therefore,

$$\begin{aligned} 0 &< 100L(L - 2.6)^2 - 23L^2 + 132L - 114 = E(L) - 3L^2 \\ &< E(L) - \frac{4(77)^3}{27(84)^2} L^2 = E(L) + L^2 \min_{0 \leq t \leq 1} t^2(84t - 77) \\ &\leq E(L) - 77L^2t^2 + 84L^2t^3. \end{aligned} \tag{2.7}$$

Thus, the expression (2.6) will be proved positive if we show that $0 < F(L) - 84L^2 = h(L)$ (say). But an easy computation shows that $h(1.6) > 0$, $h'(1.6) > 0$, and $h''(L) > 0$ for $L \geq 1.6$.

In the region $1 \leq L \leq \lambda$, we have $C(L) \geq 0$, and thus need only show that

$D(L) + B(L)t^2 + A(L)t^3 > 0$. Because of (2.5), this is immediate if $B(L) \geq 0$, so we need only consider the possibility $B(L) < 0$, in which case

$$\begin{aligned} (2.8) \qquad & D(L) + B(L)t^2 + A(L)t^3 \\ & \geq [D(L) + B(L) \max_{0 \leq t \leq 1} (t^2 - t^3/2)] + t^3[A(L) + B(L)/2] \\ & = [D(L) + B(L)/2] + t^3[A(L) + B(L)/2] = R(L) + S(L)t^3 \text{ (say)}. \end{aligned}$$

One sees easily that $R(L) > 0$ for $1 \leq L \leq 2$ and $S(L) \geq 0$ for $L \geq 1$, completing the proof.

3. Symmetric regions; the cube

The case of the q-cube (1.9) exhibits less regularity than either the simplex or ball. This is seen even in the linear case described in section 1.3, where more than k points of support may be required (as when $q = 2$ and the unique optimum design is uniform on the 4 corners); and, even more, in the quadratic case, where optimum designs can be written down explicitly almost immediately for the ball and simplex, but require at least some consideration for the cube, regarding weights assigned to the points of the 3^q array J of points with coordinates $0, 1, -1$. We shall now study this quadratic case in considerable detail. We begin by characterizing some properties of optimum quadratic designs for more general symmetric regions. (For general linear regression see Kiefer [6].)

3.1. *Quadratic regression on symmetric regions.* We introduce some of the ideas by considering, in the present paragraph, the q-cube. The fact that, when $m = 2$, the support of every optimum ξ^* is a subset of the 3^q array J, is easily seen as follows: $d(x, \xi^*)$ for any optimum ξ^* is symmetric under the group of symmetries of the cube (see discussion just below (1.5)), goes to $+\infty$ with $|x|$, and is a positive quartic on Euclidean q-space. Writing B for the subset of X where $d(x, \xi^*) = k$ (so that the support of ξ^* is contained in B), we will show that the existence of points in $B - J$ leads to a contradiction. Calling vertices, edges, and so on, the 0-, 1-, $\cdots$, skeleton of X, suppose that x' in $B - J$ lies in the r-skeleton, and hence in the relative interior of some r-cube G of that skeleton. By symmetry of B, there is another point x'' of $B - J$ which is also in the relative interior of G. The function d attains its maximum on G at x' and x'', and hence cannot be a positive quartic on q-space unless it is a constant, in which case it does not go to $+\infty$ with $|x|$.

We turn now to more general symmetric regions X to which we can apply some similar arguments.

We consider quadratic regression in q variables $x_1, \cdots, x_q$ on a symmetric region X of Euclidean q-space. The meaning of saying X is symmetric is that X is invariant under permutations $(x_1, \cdots, x_q) \to (x_{\sigma_1}, \cdots, x_{\sigma_q})$ and is invariant under sign changes $(x_1, \cdots, x_q) \to (\epsilon_1 x_1, \cdots, \epsilon_q x_q)$, $\epsilon_1 = \pm 1, \cdots, \epsilon_q = \pm 1$. The discussion just below (1.5) states that there are optimum designs which are symmetric; also, it implies that the function $d(\cdot, \xi^*)$ for any optimum design ξ^*

is a symmetric polynomial in the variables $x_1^2, \cdots, x_q^2$, of degree 2 in these variables. If we write $s = x_1^2 + \cdots + x_q^2$ and $t = x_1^4 + \cdots + x_q^4$, then the general polynomial of this type is

$$P(s, t) = as^2 + bs + c + dt. \tag{3.1}$$

The map $h: (x_1, \cdots, x_q) \to (x_1^2 + \cdots + x_q^2, x_1^4 + \cdots + x_q^4)$ maps the region X to a region X^* in the s, t plane. Clearly, $d(h^{-1}(s, t), \xi) = d^*((s, t), \xi)$ (say) is well-defined for any symmetric ξ and any (s, t) in X^*. We will be concerned with an examination of the values of $d^*(\cdot, \xi^*)$ at points of X^*, for an optimum ξ^*. We know from (1.4) and (1.6) that, throughout X,

$$d(x, \xi^*) - (q + 1)(q + 2)/2 \leq 0, \tag{3.2}$$

the equality holding at points including the support of ξ. We now show that there are two possibilities:

(i) the zeros of $d^*(\cdot, \xi) - (q + 1)(q + 2)/2$ lie entirely on the boundary of X^*;

(ii) the coefficient $d = 0$ in (3.1), so that the polynomial has the form $as^2 + bs + c$. (In this case the design is a rotatable design which can be shown to be optimum for the problem wherein X is replaced by the smallest ball centered at the origin and containing X, minus the largest open ball contained in its complement (which subtraction is vacuous if X contains the origin).

To see the validity of this assertion, suppose (s_0, t_0) is an interior point of X^* at which $P(s_0, t_0) - (q + 1)(q + 2)/2 = 0$. In view of (3.2) and continuity of the map h, (s_0, t_0) is a local maximum of the polynomial P. Therefore, the first partial derivatives vanish at (s_0, t_0), so that $d = 0$ follows. That proves the assertion.

An analysis of which of (i) and (ii) holds requires more precise knowledge of X, as we see by contrasting the cases (1.8) and (1.9). Although we already know from the first paragraph of this subsection that (i) holds for the q-cube (1.9), we shall continue our analysis for that example along the present lines, both to illustrate this method which can be applied to other symmetric regions similarly, and also because we will then use the method for cubic regression on the q-cube.

Thus, we now suppose X is given by (1.9). We will see that X^* is a closed bounded set which may be described in terms of an upper and lower boundary curve. The upper curve consists of q pieces:

$$\{t = (s - i)^2 + i, \quad i \leq s \leq i + t\}, \qquad i = 0, 1, \cdots, q - 1. \tag{3.3}$$

The lower curve may be described by the single equation

$$\{qt = s^2, 0 \leq s \leq q\}. \tag{3.4}$$

This last assertion follows at once by the Cauchy-Schwarz inequality since

$$s^2 = (x_1^2 + \cdots + x_q^2)^2 \leq q(x_1^4 + \cdots + x_q^4) = qt. \tag{3.5}$$

We observe that equality holds in (3.5) if and only if $x_1^2 = x_2^2 = \cdots = x_q^2$.

To obtain the upper boundary we suppose the value of $t = x_1^4 + \cdots + x_q^4$ is fixed and seek to minimize s. We may suppose at the start that $x_1 \geq x_2 \geq \cdots \geq x_q \geq 0$.

Consider x_q as a function of x_1 and take partial derivatives with $x_2, \cdots, x_{q-1}$ fixed. This gives $\partial x_q/\partial x_1 = -x_1^3/x_q^3$ and $\partial s/\partial x_1 = 2x_1(x_q^2 - x_1^2)/x_q^2$. We suppose here $x_q > 0$. Since the derivative is negative, we decrease s by increasing x_1 and decreasing x_q, and this preserves the ordering $x_1 \geq x_2 \geq \cdots \geq x_q \geq 0$.

Using this it may be seen that if $t = i + \delta^4$, $0 \leq \delta < 1$, is the fixed value of t, then the minimum for s is obtained by taking $x_1 = x_2 = \cdots = x_i = 1$, $x_{i+1} = \delta$, $x_{i+2} = \cdots x_q = 0$. Thus $s = i + \delta^2$ and $t = i + (s - i)^2$, as asserted in (3.3).

This argument shows even more, that the minimum value of s can be obtained from $x_1, \cdots, x_q$ if and only if $x_1 = 1, \cdots, x_i = 1$ (when $t = i + \delta^4$).

We now show, using an argument like that of the first paragraph of this subsection, that the only possible location of a zero of $d^*(\cdot, \xi^*)$ (for ξ^* optimum) on the boundary segment $\{t = i + (s - i)^2, i \leq s \leq i + 1\}$, is at an end point. The polynomial $d^*((s, i + (s - i)^2), \xi^*)$ is a quadratic in s for $0 \leq s < \infty$ which, being equal to $f(x)M^{-1}(\xi^*)f(x)'$ with $x_1 = x_2 = \cdots = x_i = 1$, $x_{i+1} = s^{1/2}$, $x_{i+2} = \cdots = x_q = 0$, is nonnegative for $0 \leq s < \infty$ and goes to infinity with s, so that it cannot have a local maximum over the interval $i \leq s \leq i + 1$ at an interior point of the latter.

Finally, using the same type of argument, we show that no zero of $d^*(\cdot, \xi^*) - k$ can occur interior to the segment $\{qt = s^2, 0 \leq s \leq q\}$. This is so because $d^*((s, s^2/q), \xi^*)$ is a quadratic in s for $0 < s < \infty$ which, being equal to $f(x)M^{-1}(\xi^*)f(x)'$ with $x_1 = x_2 = \cdots = x_q = (s/q)^{1/2}$, goes to infinity with s, and can thus not attain its maximum over $0 \leq s \leq q$ at an interior point of the latter.

Our discussion has not yet excluded the possibility that the optimal design is rotatable, that is, that $d^*((s, t), \xi^*)$ has the form $as^2 + bs + c$ for an optimum ξ^*. Were this the case, then the design would be optimum for X replaced by the ball $K = \{x: \sum_1^q x_i^2 \leq q\}$, since the argument of the previous paragraph shows that if the optimum ξ^* is rotatable, then $d(x, \xi^*)$ takes on its maximum value $(q + 1)(q + 2)/2$ at the point $(1, 1, \cdots, 1)$ satisfying $\sum_1^q x_i^2 = q$. The moment matrix $M(\xi)$ for an optimal design ξ on K is uniquely determined and is known (Kiefer [6]) to put mass at $s = 0$ and on $s = q$, the moments of the conditional distribution on $s = q$ being those of the uniform measure on this surface.

But, in our problem, the design ξ must be concentrated in the cube (1.9), and the only points in common between the cube and the shell $s = q$ are the corners $(\pm 1, \cdots, \pm 1)$. It is easily seen that every symmetric ξ which is concentrated on the corners and at the origin makes $\int x_1^4 \xi(dx) = \int x_1^2 x_2^2 \xi(dx)$, which is not the case for the optimum rotatable design on the ball. Hence the optimal design for quadratic regression on the q-cube cannot be rotatable.

The discussion substantiates what was already known from simpler calculations in this special case, as indicated earlier. We now bring these ideas to bear on the problem of cubic regression on the cube. We shall return in sections 3.3–3.5 to quadratic regression and shall consider at length there the possible supporting sets for optimum designs.

3.2. *Cubic regression on the q-cube.* We first treat the case $q = 2$, for the sake of simplicity and explicitness of numerical results. The function $d(\cdot, \xi^*)$ for an optimum ξ^* is now a nonnegative polynomial of degree 6 in the variables x_1, x_2, having the following properties. From (1.4) and (1.6),

$$\bar{d}(\xi^*) \leq 10; \tag{3.6}$$

and d is symmetric in x_1, x_2 and invariant under sign changes.

It follows that if we write $t = x_1^4 + x_2^4$ and $s = x_1^2 + x_2^2$ as before, then we can define $d^*(\cdot, \xi)$ for symmetric ξ as in the paragraph below (3.1). This will now be a polynomial of degree 3 in s, t, of the form

$$P(s, t) = as^3 + bs^2 + cs + d + est + ft. \tag{3.7}$$

The domain X^*, from (3.3) and (3.4) with $q = 2$, is the closed bounded set whose boundary consists of the three curves

$$\begin{aligned} &\{s^2 = t, 0 \leq s \leq 1\}, \\ &\{(s-1)^2 + 1 = t, 1 \leq s \leq 2\}, \\ &\{s^2 = st, 0 \leq s \leq 2\}. \end{aligned} \tag{3.8}$$

We consider first the implication of assuming that for some (s_0, t_0) interior to this region $d((s_0, t_0), \xi^*) = 10$. This would be a local maximum interior to the domain X^*, so that $P(\sigma + s_0, \tau + t_0) = a\sigma^3 + b'\sigma^2 + c'\sigma + d' + e\sigma\tau + f\tau$ (say), defined in a neighborhood of $(\sigma, \tau) = (0, 0)$, would have a local maximum at $(0, 0)$. We would therefore have $e = f = 0$, and $P(s, t)$ would be a function only of s, which is the definition of ξ^* being rotatable. We also note that in this case we would have

$$P(s, t) = (as + b'')(s - s_0)^2 + 10 \tag{3.9}$$

with $as + b'' \leq 0$ for $0 \leq s \leq 2$ and with $a > 0$ (since $P \to \infty$ as $s \to \infty$).

Thus, if the design is not rotatable, then the polynomial $d^*(\cdot, \xi^*)$ can vanish only on the boundary curves of (3.8). Substitution of any one of the three relations of (3.8) for t into $P(s, t) - 10$ gives a cubic in s which does not change sign. Hence, any root s interior to the interval determined by the substitution would require the root to be a double root. Therefore, each of the three sections of boundary in the s, t plane can have at most two points at which the polynomial $P - 10$ vanishes. Since the boundary of the square X is mapped into the curve $\{(s-1)^2 + 1 = t, 1 \leq s \leq 2\}$, it follows from the final paragraph of section 1.2 that this curve contains at least one point at which $P - 10$ vanishes.

We now eliminate the possibility of a rotatable design being optimum. The theory of optimum designs for polynomial regression on the ball has been developed by Kiefer [6] and in section 4 of the present paper. An optimum design ξ^* for the square, if rotatable, would have a $d^*(\cdot, \xi^*)$, given by (3.9), attaining its maximum on the square at the corners ($s = t = 2$) and satisfying $d^*((s, t), \xi^*) \leq 10$ on the image under h of the ball K of radius $2^{1/2}$. (This image is bounded by the curves $\{s^2 = 2t, 0 \leq s \leq 2\}$, $\{s^2 = t, 0 \leq s \leq 2\}$, and

$\{s = 2, 2 \leq t \leq 4\}$.) Hence ξ^* would be optimum for the problem of cubic regression on $K = \{x: x_1^2 + x_2^2 \leq 2\}$, with $d(x, \xi^*)$ attaining its maximum over K on the two circles $s = 2$ and $s = s_0$. Thus, $s_0 = 2\rho^2$, where ρ and 1 are the radii of circles where $d(x, \xi') = 10$, where ξ' is optimum for cubic regression when X is the unit ball of (1.8), to be discussed further in section 4. Although this ξ' is not unique, its moments up to those of order 4 (that is, the elements of $M(\xi')$), and the total masses β and $1 - \beta$ which it assigns to the circles of radii ρ and 1, respectively, are unique. (For the first of these facts, see just above (1.5); for the second, replace ξ' by the optimum design ξ'' defined by $\xi''(A) = \int \xi'(gA)\nu(dg)$, where ν is the invariant probability measure on the orthogonal group in two dimensions as in [6], and use the uniqueness of the masses in such a ξ'', proved in [6].) Since $\int_0^{2\pi} (2\pi)^{-1} \cos^4 \theta \, d\theta = \frac{3}{8}$, we would thus obtain

$$\int x_1^4 \xi^*(dx) = 4 \int x_1^4 \xi'(dx) = 4 \left(\tfrac{3}{8}\right) [(1 - \beta) + \rho^4 \beta]. \tag{3.10}$$

Since ξ^* is supported within the square (1.9), we also have $\int x_1^4 \xi^*(dx) \leq 1$. Hence, (3.10) yields

$$\frac{3(1 - \beta)}{2} \leq \left(\tfrac{3}{2}\right) [(1 - \beta) + \rho^4 \beta] \leq 1, \tag{3.11}$$

or $\beta \geq \frac{1}{3}$. But (from table 4.1 of section 4) $\beta = .3077$. We conclude that the optimum design cannot be rotatable.

The following maximization problem was solved numerically. Put mass $p_1/4$ at each of $(1, 1)$, $(1, -1)$, $(-1, 1)$ and $(-1, -1)$. Put mass $p_2/8$ at each of the eight points $(\pm 1, \pm a)$ and $(\pm a, \pm 1)$, $0 < a < 1$. Put mass $(1 - p_1 - p_2)/4$ at each of the four points $(\pm b, \pm b)$, $0 < b < 1$. (In each of the above, the two $\pm$ signs act independently.) Our earlier discussion shows a design of this form to be a candidate for being optimum (although we did not yet eliminate certain other forms). The determinant of $M(\xi)$ was maximized on the Cornell CDC 1604 as a function of the four parameters involved, giving

$$\begin{aligned} a &= 0.35880, \\ b &= 0.48000, \\ p_1 &= 0.36770, \\ p_2 &= 0.46100. \end{aligned} \tag{3.12}$$

For this design ξ, the quantity $\bar{d}(\xi)$ was computed numerically and was found to be ≤ 10 to five decimal places.

We now leave the case $q = 2$ to discuss cubic regression on the cube (1.9) for general $q \geq 3$, where an analysis similar to the one just given for $q = 2$ may again be carried out. If ξ is symmetric, $d(x, \xi)$ is now a sixth degree symmetric polynomial in $x_1, \cdots, x_q$, in which any monomial term involving an odd exponent has zero coefficient. We now need three symmetric functions,

$$
\begin{aligned}
s &= x_1^2 + \cdots + x_q^2, \\
t &= x_1^4 + \cdots + x_q^4, \\
u &= x_1^6 + \cdots + x_q^6,
\end{aligned}
\tag{3.13}
$$

and define $h: X \to X^*$ by $h(x) = (s, t, u)$. (When $q = 2$, we do not need u since $2u = 3st - s^3$ in that case.) Again $d^*(\cdot, \xi)$ on X^* is well-defined for symmetric ξ by $d^*(h(x), \xi) = d(x, \xi)$, and now has the form

$$P(s, t, u) = as^3 + bs^2 + cs + d + est + ft + gu. \tag{3.14}$$

If this polynomial has a local maximum in the interior of X^*, then one may show $e = f = g = 0$ as before, and therefore one may conclude that the design is rotatable. (In fact, this conclusion clearly holds if (1.9) is replaced by an arbitrary compact symmetric q-dimensional set.)

We now extend the argument we used when $q = 2$, to conclude again that an optimum design ξ^* cannot be rotatable. Such a ξ^* would, by the same argument as before, be optimum for cubic regression on $K = \{x: x_1^2 + \cdots + x_q^2 \leq q\}$. Let β and $1 - \beta$ again denote the total mass assigned to the spheres $\{\sum_1^q x_i^2 = \rho\}$ and $\{\sum_1^q x_i^2 = 1\}$, by each optimum design for cubic regression on the unit ball (1.8). The integral of x_1^4, with respect to the uniform probability measure on $\{\sum_1^q x_i^2 = 1\}$, is now $3/q(q + 2)$. Also, just as before, $\int x^4 \xi^*(dx) \leq 1$. Thus, the analogue of (3.10) and the second inequality of (3.11) is that, if ξ^* is an optimum design for K, then

$$q^2[3/q(q + 2)][(1 - \beta) + \beta\rho^4] = \int x_1^4 \xi^*(dx) \leq 1. \tag{3.15}$$

If one writes down equations from which the parameters of an optimum rotatable design on (1.8) may be calculated, then one obtains (see section 4, equation (4.5))

$$
\begin{aligned}
&\frac{(q + 3)(q + 2)(q + 1)}{6} \\
&\qquad = \frac{q + 1}{1 - \beta} + \frac{(q - 1)(q + 2)}{2((1 - \beta) + \beta\rho^4)} + \frac{(q + 4)q(q - 1)}{6((1 - \beta) + \beta\rho^6)}.
\end{aligned}
\tag{3.16}
$$

Since $\rho < 1$, we may replace ρ^6 by ρ^4 and also drop the first term on the right in (3.16), and may then divide both sides by $(q + 1)/6$, obtaining

$$q^2 + 5q + 6 > [q^2 + 5q - 6]/[(1 - \beta) + \beta\rho^4]. \tag{3.17}$$

Substituting this inequality for $1 - \beta + \beta\rho^4$ into (3.15) yields

$$
\begin{aligned}
0 &> 3q[q^2 + 5q - 6] - (q + 2)[q^2 + 5q + 6] \\
&= 2q^3 + 8q^2 - 34q - 12.
\end{aligned}
\tag{3.18}
$$

The last polynomial is easily seen to be positive for $q \geq 3$. We have thus proved theorem 3.1.

THEOREM 3.1 *For cubic regression on the q-cube* (1.9) *with* $q \geq 2$, *an optimum design cannot be rotatable.*

Thus, the problem reduces to a study of the nature of the boundary of X^* in the (s, t, u)-space and of the solution of the appropriate maximization problem. We do not attempt to do this in the present paper.

3.3. *Optimum symmetric designs for quadratic regression on the q-cube.* We have already given a short proof in the first paragraph of section 3.1, that all optimum designs for $m = 2$ on the q-cube (1.9) are supported by a subset of the 3^q-array J. For $j = 0, 1, \cdots, q$, let J^j be the subset of J consisting of those $2^{q-j}\binom{q}{j}$ points with exactly j coordinates equal to zero. (Thus, J^j consists of the midpoints of all j-dimensional faces of X.) Kiefer [7] obtained optimum designs supported by $J^0 \cup J^1 \cup J^2$ when $q \leq 5$, showed that this set could not support an optimum design when $q \geq 6$, and described ([7], footnote 5) the method, obtained with Farrell, for obtaining optimum designs for each q on the union of three J^j's and certain subsets of such a union by solving (3.22) below. This joint work is the subject of the present subsection. Subsequently, Kono [9], citing this description in [7], also showed that optimum symmetric designs for each q can only be supported by a subset of J, and carried out detailed calculations for an optimum design on the union of three J^j's when $q \leq 9$, obtaining optimum symmetric designs on $J^0 \cup J^1 \cup J^q$, again by solving equations (3.22) below, for that choice of the J^j's.

We shall first characterize those sets of the form

$$J(j_1, j_2, j_3) = \bigcup_{i=1}^{3} J^{j_i} \tag{3.19}$$

which can support a symmetric optimum design. We shall take $j_1 < j_2 < j_3$. (It can be seen that two J^j's cannot suffice when $q \geq 2$, by noting that no α_{j_i} can be 0 in the demonstration which follows.) Such a design assigns probability $\alpha_{j_i}/2^{q-j_i}\binom{q}{j_i} > 0$ to each point of $J^{j_i}(i = 1, 2, 3)$, where $\sum_1^3 \alpha_{j_i} = 1$. The pertinent moments of such a design are computed, as in (4.1) of [7], to be

$$\begin{aligned} u(\xi^*) &= \int x_1^2\xi^*(dx) = \int x_1^4\xi^*(dx) = \sum_1^3 \alpha_{j_i}(q - j_i)/q, \\ v(\xi^*) &= \int x_1^2x_2^2\xi^*(dx) = \sum_1^3 \alpha_{j_i}(q - j_i)(q - j_i - 1)/q(q - 1). \end{aligned} \tag{3.20}$$

Write as in (4.5) of [7],

$$\begin{aligned} U_q &= \frac{(q + 3)}{4(q + 1)(q + 2)^2}\{(2q^2 + 3q + 7) \\ &\qquad + (q - 1)[4q^2 + 12q + 17]^{1/2}\}, \\ V_q &= \frac{(q + 3)}{8(q + 2)^3(q + 1)}\{(4q^3 + 8q^2 + 11q - 5) \\ &\qquad + (2q^2 + q + 3)[4q^2 + 12q + 17]^{1/2}\}. \end{aligned} \tag{3.21}$$

One computes $M^{-1}(\xi^*)$, as in (4.3) of [7], and then observes, exactly as in [7], that $\bar{d}(\xi^*) = k$ if and only if

$$u(\xi^*) = U_q, \qquad v(\xi^*) = V_q. \tag{3.22}$$

To solve these equations for the j_i and α_{j_i}, we may think of plotting in the (u, v)-plane, for fixed $q \geq 2$, the $q + 1$ points

$$(u_j, v_j) = ((q - j)/q, (q - j)(q - j - 1)/q(q - 1)), \qquad 0 \leq j \leq q, \tag{3.23}$$

and the point (U_q, V_q). Then clearly (3.22) is satisfied for nonnegative j_i if and only if (U_q, V_q) lies in the triangle with vertices (u_{j_i}, v_{j_i}). Even though we shall not use any such geometric considerations in the demonstration which follows, they help in understanding the results, and also in understanding what is involved in considering unions of more than three J^i's, which we shall forego. (We also remark to the reader that the idea of the demonstration which follows, without such tedious computational details, can be obtained by replacing U_q and V_q by their asymptotic values $2 - q^{-1} + 3q^{-2}/2$ and $1 - 2q^{-1} + 5q^{-2}$, and following through the argument for "large q".)

We first note, replacing $[4q^2 + 12q + 17]^{1/2}$ by the smaller value $(2q + 3)$, that it is easy to verify that $U_q > (q - 1)/q \geq u_j, j > 0$, from which we conclude that $j_1 = 0$. Substituting this fact into (3.22) (and using $\sum_1^3 \alpha_{j_i} = 1$), we obtain

$$\begin{aligned} \alpha_{j_2} &= [q/j_2(j_3 - j_2)]\{-(q - 1)V_q + (2q - j_3 - 1)U_q - (q - j_3)\}, \\ \alpha_{j_3} &= [q/j_3(j_3 - j_2)]\{(q - 1)V_q - (2q - j_2 - 1)U_q + (q - j_2)\}. \end{aligned} \tag{3.24}$$

For fixed q, the expression $\{(q - 1)V_q - (2q - j - 1)U_q + (q - j)\} = F_q(j)$ (say) is linear and (since $U_q < 1$) decreasing in j. We shall show in the next two paragraphs that

$$F_q(2) > 0 > F_q(3) \qquad \text{for} \quad q > 5. \tag{3.25}$$

It follows then from (3.24) that (3.22) can be satisfied for $q > 5$ (with positive α_{j_i}'s) if and only if

$$0 < j_2 < 3 \leq j_3. \tag{3.26}$$

The first inequality of (3.25) can be written as

$$q - 2 > (2q - 3)U_q - (q - 1)V_q, \tag{3.27}$$

or

$$\begin{aligned} 8(q - 2)(q + 2)^3(q + 1)/(q + 3) > 4q^4 + 12q^3 + 7q^2 - 6q \\ - 89 + (2q^3 - q^2 - 16q + 15)[4q^2 + 12q + 17]^{1/2}. \end{aligned} \tag{3.28}$$

A direct computation shows that the left side of (3.28) equals

$$\begin{aligned} 8\{q^4 + 2q^3 - 2q^2 - 10q - 4 + (2q - 4)/(q + 3)\} \\ > 8q^4 + 16q^3 - 16q^2 - 80q - 32. \end{aligned} \tag{3.29}$$

Using the fact that $[4q^2 + 12q + 17]^{1/2} < 2q + 3 + 4/(2q + 3)$ and that $4(2q^3 - q^2 - 16q + 15)/(2q + 3) < 2q(2q - 3)$, we obtain that the right side of (3.28) is less than $8q^4 + 16q^3 - 24q^2 - 30q - 44$. This last is less than the right side of (3.29) by $(q - 6)(8q + 2)$, proving (3.28) (and thus (3.27)) for $q \geq 6$.

The second inequality of (3.25) can be written as

$$q - 3 < (2q - 4)U_q - (q - 1)V_q, \tag{3.30}$$

or

$$8(q - 3)(q + 2)^3(q + 1)/(q + 3) < 4q^4 + 8q^3 - 7q^2 - 32q - 117 + (2q^3 - 3q^2 - 18q + 19)[4q^2 + 12q + 17]^{1/2}. \tag{3.31}$$

The left side of (3.31) is

$$8\{q^4 + q^3 - 6q^2 - 16q - 4 - 12/(q + 3)\} < 8q^4 + 8q^3 - 48q^2 - 128q - 32. \tag{3.32}$$

Using the fact that $[4q^2 + 12q + 17] > 2q + 3 + 4/(2q + 3) - 16/(2q + 3)^3$ and that $[4/(2q + 3) - 16/(2q + 3)^3](2q^3 - 3q^2 - 18q + 19) > 4q^2 - 12q - 22$, we obtain that the right side of (3.31) is greater than $8q^4 + 8q^3 - 48q^2 - 60q - 82$. The latter is clearly greater than the right side of (3.32), proving (3.31) and thus (3.30).

In the same manner that (3.26) was proved (or by direct calculation in the few cases $q < 6$), one can show that (3.26) is replaced by $0 < j_1 < 2 \leq j_2$ for $2 \leq q \leq 6$. To summarize, then, we have the following theorem.

THEOREM 3.2. *The set $J(j_1, j_2, j_3)$ of (3.19) supports a symmetric optimum design for quadratic regression on the q-cube, if and only if*

$$\begin{aligned} &j_1 = 0, \quad j_2 = 1, \quad 2 \leq j_3 \leq q, \quad &&\text{when } 2 \leq q \leq 5, \\ &j_1 = 0, \quad j_2 = 1 \text{ or } 2, \quad 3 \leq j_3 \leq q, \quad &&\text{when } q \geq 6. \end{aligned} \tag{3.33}$$

We remark that, among the sets $J(0, j_2, j_3)$ permitted by (3.33), the set $J(0, 1, q)$ consisting of vertices, midpoints of edges, and center, has the smallest number of points $(2^q + q2^{q-1} + 1)$ of any optimum symmetric design. In view of (1.12), such designs are quite unsatisfactory for large q, and in the remainder of section 3 we shall therefore seek asymmetric designs on fewer points.

The weights α_{j_i} for any optimum symmetric design on a set (3.19) permitted by (3.33) may be obtained from (3.24) and $\alpha_0 = 1 - \alpha_{j_2} - \alpha_{j_3}$. For $j_2 = 1$, $j_3 = 2$, and $q \leq 5$, $\alpha_{j_i}/2^{q-j_i}\binom{q}{j_i}$ is tabled in [7]; for $j_2 = 1$, $j_3 = q \leq 7$, the α_j are tabled in [9].

3.4. *Bounds on N for quadratic regression on the q-cube.* We now apply the considerations of section 1.3 regarding the minimum number N of points needed to support an optimum design on the q-cube when $m = 2$. We first note

THEOREM 3.3. *The optimum design for quadratic regression on the q-cube is unique if and only if $q \leq 2$, in which case the support is J.*

PROOF. The lack of uniqueness when $q \geq 3$ follows from (3.33). The uniqueness when $q = 1$ or 2 (the former of which is well known) can be proved by using (1.16) with $U = J$; the matrix $\{\phi_t(x), x \in J, 1 \leq t \leq p\}$ is easily seen to have rank 3 or 9 in these two cases. (We recall that the subscript value $t = 0$ need not be included in polynomial regression.)

We next improve the upper bound $\binom{q+4}{4}$ of (1.12) by use of (1.14). We shall take $U = W = J(0, 1, q)$ in the calculation which follows. For $x \in J(0, 1, q)$ (in fact, for $x \in J$), the relations

$$\begin{aligned} x_i &= x_i^3, \\ x_i^2 &= x_i^4, \\ x_i x_j &= x_i x_j^3, \qquad i \neq j \end{aligned} \tag{3.34}$$

for $1 \leq i, j \leq q$ are satisfied. These are $q^2 + q$ linearly independent relations among the $\phi_t (1 \leq t \leq p)$. Among the set of ϕ_t which remain after deleting those on the right side of (3.34), the relations

$$\begin{aligned} (1 - x_j^2)(x_r^2 - x_s^2) &= 0, \\ (x_i^2 - x_j^2)(x_r^2 - x_s^2) &= 0, \end{aligned} \tag{3.35}$$

are satisfied on $J(0, 1, q)$ with all subscripts unequal and between 1 and q, inclusive. (Equalities (3.35) are vacuous if $q < 3$.) This is so because either all $x_i^2 = 0$, or else at most one $x_i^2 = 0$. The relations (3.35) among the ϕ_t are not linearly independent when $q \geq 3$, so we must find the dimension of the vector space spanned by the ϕ_t. To this end, we write $L = q + 1$, $y_i = x_i^2$ for $1 \leq i \leq q$, and $y_{L+1} = 1$.

For $L \geq 4$, let Q be the vector space over the reals of all linear combinations of the polynomials $(y_i - y_j)(y_r - y_s)$ with i, j, r, s distinct integers between 1 and L, inclusive (a subspace of the quadratic polynomials in L variables). We shall show the following lemma.

LEMMA 3.4. *For $L \geq 4$, we have*

$$\dim Q = L(L - 3)/2. \tag{3.36}$$

PROOF. All subscripts in the proof which follows are to be reduced mod L. We first show that the $L(L - 3)/2$ special polynomials

$$(y_j - y_{j+1})(y_{j+i+1} - y_{j+i+2}), \tag{3.37}$$

with all subscripts distinct, span Q. (Note, for example, that $j = L$ is permitted.) We must show that any polynomial $(y_i - y_j)(y_r - y_s)$ is a linear combination of these special polynomials. There are two cases to which any other can be reduced by symmetry.

Case 1 $(i < j < r < s)$. Then

$$(y_i - y_j)(y_r - y_s) = \sum_{0 \leq u < j-i, 0 \leq v < s-r} (y_{i+u} - y_{i+u+1})(y_{r+v} - y_{r+v+1}). \tag{3.38}$$

Case 2 $(i < r < j < s)$. Use the identity

$$(y_i - y_j)(y_r - y_s) = (y_i - y_r)(y_j - y_s) - (y_r - y_j)(y_s - y_i) \tag{3.39}$$

to reduce to case 1.

To conclude the proof of (3.36), we need only show that the special polynomials (3.37) are linearly independent. To this end, we obtain an appropriate ordering of the special polynomials, say $\{g_\alpha, 1 \leq \alpha \leq L(L-3)/2\}$, and of the functions $y_i y_j (i \neq j)$, say $\{h_\beta, 1 \leq \beta \leq L(L-1)/2\}$; then writing $g_\alpha = \sum_\beta c_{\alpha\beta} h_\beta$, we show that $c_{\alpha\alpha} \neq 0$ and $c_{\alpha\beta} = 0$ for $\alpha > \beta$, which proves the desired result. The g_α are, in order, the special polynomials of (3.37) with $i = 1$ and $j = 0, 1, \cdots, L-1$; then, with $i = 2$ and $j = 0, 1, \cdots, L-1; \cdots, i = \{$greatest integer $\leq (L-3)/2\}$ and $j = 0, 1, \cdots, L-1$; if L is even, there are then $L/2$ additional special polynomials with $i = (L-2)/2$ and $j = 0, 1, \cdots, (L-2)/2$. Note that those functions in the ith collection of L polynomials (or of $L/2$ if L is even and $i = (L-2)/2$) have i as the minimum distance between subscripts γ, δ which appear in any term $y_\gamma y_\delta$ entering with nonzero coefficient in the special polynomial. Moreover, for fixed i, these $y_\gamma y_{\gamma+i}$ appear in the order $\gamma = 1, 2, \cdots, L$. Thus, when we order the h_β as $y_j y_{j+1}$ for $j = 1, 2, \cdots, L$, and then $y_j y_{j+2}$ for $j = 1, \cdots, L$, and so on, we see at once that the $c_{\alpha\beta}$ have the desired property. This completes the proof of the lemma.

Putting $L = q + 1$ in (3.36), we conclude that the ϕ_t on the left side of (3.35) span a real vector space of dimension $(q+1)(q-2)/2$ (which is also correct if $q = 2$). Adding this to the number $q^2 + q$ of restrictions (3.34), which are independent of (3.36), and subtracting the result from $p = \binom{q+4}{4}$ and using (1.14) with $\{0 \leq t \leq p\}$ replaced by $\{1 \leq t \leq p\}$, we obtain theorem 3.5.

THEOREM 3.5. *For quadratic regression on the q-cube,*

$$N \leq (q+1)(q^3 + 9q^2 - 10q + 48)/24. \tag{3.40}$$

3.5. *The use of orthogonal arrays to reduce the number of points of support.* We have already mentioned in section 1 how orthogonal arrays of strength 2 are used classically to reduce the number of points of support for an optimum design for linear regression on the q-cube. Similar techniques can be employed in other settings, as we now illustrate for quadratic regression on the q-cube. We shall consider the following particular scheme of application.

Suppose, for each positive integer r, that we can find a subset A_r of the 2^r corners of the r-cube ((1.9) with $q = r$), such that the uniform probability measure on A_r has the same moments of order ≤ 4 as the uniform probability measure on the 2^r corners. Suppose A_r has n_r points. Then, suppose we replace J^0 in $J(0, 1, q)$ by A_q (with α_0/n_q probability per point); replace the 2^{q-1} points of J^1 with $x_i = 0$ by the n_{q-1} points of the form

$$x_i = 0, \qquad (x_i, \cdots, x_{i-1}, x_{i+1}, \cdots, x_q) \in A_{q-1},$$

for $1 \leq i \leq q$ (with probability α_1/qn_{q-1} per point); retain J^q, with probability α_q. Since only moments of order ≤ 4 are present in $M(\xi)$, and because of the way in which zero coordinate values enter into the replacement of J^1, we obtain a design with the same M as the optimum symmetric design on $J(0, 1, q)$, and which is therefore also optimum. It is supported by

$$n_q + qn_{q-1} + 1 \tag{3.41}$$

points.

The classical construction of such an A_r is in terms of an *orthogonal array of strength 4 with 2 levels*, that is, an $n_r \times r$ matrix T_r with entries ± 1 such that every 4-row submatrix has the property that each of the 16 possible 4-vectors with entries ± 1 appears equally often in the submatrix. We then consider the n_r columns of the matrix T_r as the points of A_r and clearly obtain the required moment properties. The reader is referred to such references as [1] for detailed discussion of orthogonal arrays.

Orthogonal arrays of strength ≥ 3 have been considered extensively by Rao, Bose, Bush, Seiden, and others. The principal method of construction is geometric. A set C of r points in the finite projective space $PG(d-1, 2)$ of dimension $d-1$, no 4 of which lie on the same 2-dimensional flat, yields a T_r with $n_r = 2^d$; for, writing B for the $r \times d$ matrix whose rows are the points of C (each of the d coordinates of such a point being an element of the Galois field $GF(2)$), and writing D for the $d \times 2^d$ matrix whose columns are the different d-vectors with coordinate values in $GF(2)$, one sees easily that BD has the required properties of T_r, except that the elements ± 1 of T_r are replaced by 0, 1 (of $GF(2)$) in BD. (It is not always known when this geometric construction yields the maximum r for given d.)

For fixed r, Rao's lower bound on n_r, usually given in geometric terms, can be obtained for general orthogonal arrays $T_n = \{t_{ij}, 1 \leq i \leq r, 1 \leq j \leq n_r\}$ of strength 4 with elements ± 1, as follows. Let τ_0 be the row vector of n_r 1's; let τ_i be the i-th row of T_r, and for $1 \leq i < i' \leq r$ let $\tau_{i,i'} = (t_{i1}t_{i'1}, \cdots, t_{in_r}t_{i'n_r})$. The $1 + r + \binom{r}{2}$ vectors $\tau_0, \tau_1, \cdots, \tau_r, \tau_{12}, \cdots, \tau_{(r-1)r}$ are easily shown to be orthogonal, because of the properties of T_r. Hence,

$$n_r \geq (r^2 + r + 2)/2. \tag{3.42}$$

In the other direction, it is simple to give a geometric construction which yields an orthogonal array of strength 4 satisfying

$$n_r = \text{largest number } \ 2^d \ \text{ which is } \leq 1 + r + \binom{r}{2} + \binom{r}{3}. \tag{3.43}$$

For, in $PG(d-1, 2)$, if we have chosen j points, no 4 of which are coplanar, there are $\binom{j}{2}$ pairs of points each of which determines a line with one point outside the pair, and $\binom{j}{3}$ triples of points, each of which determines a plane with one point not on the lines just mentioned. Thus, as long as $j + \binom{j}{2} + \binom{j}{3} < 2^d - 1$ (equal to the number of points in $PG(d-1, 2)$), there remains a $(j+1)st$ point which can be chosen without destroying noncoplanarity. Continuing in this way, we can obtain r points, where r is the smallest integer for which $r + \binom{r}{2} + \binom{r}{3} \geq 2^d - 1$. This yields (3.43).

The reader should have no trouble in writing down analogues of (3.42) and (3.43) for other Galois fields, and, in fact, for arrays of different strength.

When q is large, the use of (3.41) with (3.43) yields an optimum design on $\leq q^4(1 + o(1))/3$ points. This is $0(q^4)$ like p or the bound (3.40), but these last are both $q^4(1 + o(1))/24$. Thus, we do not know whether or not the orthogonal array approach can, for large q, yield a design with no more points than (3.40), let alone whether the order r^3 of (3.43) rather than the order r^2 of (3.42) (or neither) is the best possible as $r \to \infty$. We do know that there are some small values of q for which the method of using orthogonal arrays cannot yield a value of (3.41) which is less than (3.40) or even p. This is a consequence of the fact ([10], [12]) that the minimum possible values of n_r for $r = 4, 5, 6, 7, 8$ are known to be 16, 32, 32, 64, 64, so that the numbers listed in the last column of table I below, and which were obtained by using these values in (3.41),

TABLE I

q	k	p	(3.40)	Points in $J(0, 1, q)$	Achievable Using (3.41)
2	6	15	9	9	9
3	10	35	21	21	21
4	15	70	45	49	49
5	21	126	87	113	113
6	28	210	154	257	225
7	36	330	254	577	289
8	45	495	396	1281	577
9	55	715	590	2817	705
10	66	1001	847	6145	1409
11	78	1365	1179	13313	1537
12	91	1820	1599	28673	1793
13	105	2380	2121	61441	3585
16	153	4845	4454	589825	4353
17	171	5985	5544	1245185	4608
Asymptotic Value	$q^2/2$	$q^4/24$	$q^4/24$	$q2^{q-1}$	$\leq q^4/3$

cannot be improved upon by using orthogonal arrays for $q \leq 8$. We have also used the values $n_r = 128$ for $9 \leq r \leq 11$ and $n_r = 256$ for $12 \leq r \leq 17$ in this table. These are the best values obtainable geometrically [12], but it is not yet known whether a nongeometric construction can yield better orthogonal arrays in these cases. (For values like $q = 10$ or 13, where $n_{q-1} > n_{q-2}$, the number obtained from (3.41) is at its worst compared with p or (3.40); similarly, for $q = 8$, 9, 11, 12, and 17, the comparison is more favorable.)

We are indebted to Professor Esther Seiden for several communications concerning the construction of these orthogonal arrays of strength 4.

In view of the unattainability of p or (3.40) for some values q by using the method of this subsection, it is clear that further study is needed of designs which have less symmetry. For example, by considering nonuniform measures on smaller sets than A_r, and perhaps subsets of more than three J^i's, one should be able to do considerably better. Perhaps one can even reduce the number of points required from $0(q^4)$ to a smaller order such as $0(q^3)$. One other obvious attempt to obtain $0(q^3)$ is to seek an optimum design with equal mass on each point of J, plus additional masses on J^0 and J^q, and thus to replace the q arrays used with J^1 in (3.41) by an orthogonal array of strength 4 with three levels, which is used in place of J (the n_q and 1 being present in (3.41) as before). The analogue of (3.43) for $GF(3)$ shows that this three-level array again requires only $0(q^3)$ points, so that (3.41) would yield $0(q^3)$. Unfortunately, one cannot choose positive probabilities on J^0, J, and J^q so as to satisfy the analogue of (3.22) for large q.

4. The q-ball

We now suppose X to be the unit q-ball (1.8). For regression of degree m, a rough characterization of optimum designs was given by Kiefer [6]): every optimum ξ assigns measure one to $(m + 1)/2$ spherical shells centered at 0, where one of these shells is the boundary of X and where 0 counts as $\frac{1}{2}$ shell. Some weighted mixture of uniform measures on these shells is optimum (although other measures with the same first $2m$ moments are also optimum). The weights and radii of shells are hard to compute for $m > 2$; when $m = 2$, measure $2/(q + 1)(q + 2)$ is assigned to the origin and the remainder is assigned to the boundary of X.

Two problems of interest here are (1) to obtain at least approximate information on the radii and weights when $m > 2$, and (2) to obtain discrete measures on the shells supported by as few points as possible. In most of the remaining paragraphs of this section we shall indicate the type of treatment of problem (1) which is possible for $m > 2$, considering here the example $m = 3$. Problem (2) entails considerations related to those of section 3 and also to the extensive literature on the construction of rotatable designs. It differs from the latter in its specification of the radii and weights and in its allowing of unequal masses on points which may not be symmetrically spaced. The implementation of the resulting optimum designs of the approximate theory for specified sample sizes by discrete designs which approximate them, will yield nonrotatable designs which can be expected to involve fewer distinct points and to have better performance characteristics than the rotatable designs which are usually used. The payment for this in the form of a design matrix which is harder to invert may be worthwhile with modern computing equipment.

An optimum design ξ^* in the case $m = 3$ on the q-ball can be described in terms of two parameters: measure β is spread uniformly on a sphere of radius $\rho < 1$, and measure $1 - \beta$ is assigned to the unit sphere (equal to the boundary of X). For such a design, $d(x, \xi^*)$ depends only on $r^2 = \sum x_i^2$, say $d(x, \xi^*) = d^*(r, \xi^*)$. The optimum ρ and β can be found either by solving the two equations $d^*(\rho, \xi^*) = \binom{q+1}{3}$ $(=k)$ and $\partial d^*(r, \xi^*)/\partial r|_{r=\rho} = 0$ (see [6]), or else by maximizing $\det M(\xi)$ with respect to ρ and β. We shall exhibit the second method.

Grouping the functions into four sets as $\{1, x_1^2, \cdots, x_k^2\}$, $\{x_i, x_i^3, x_i x_j^2; i \neq j\}$,

$\{x_i x_j; i < j\}$, $\{x_h x_i x_j; h < i < j\}$, one sees that the product of two elements from different sets has zero integral. Thus, for ξ of the specified form, det $M(\xi)$ can be evaluated as the product of four determinants; one obtains, with C_q denoting a constant depending on q,

$$\text{(4.1)} \qquad \log \det M(\xi) = C_q + 2q \log \rho + (q + 1) \log [\beta(1 - \beta)(1 - \rho^2)^2] + \frac{(q + 2)(q - 1)}{2} \log [(1 - \beta) + \beta\rho^4] + \frac{(q + 4)q(q - 1)}{6} \log [(1 - \beta) + \beta\rho^6].$$

The two equations obtained by setting equal to zero the derivatives with respect to each of ρ and β, are not very manageable analytically. (This is also true of the equations obtained by the other approach mentioned in the previous paragraph.) These equations, however, can be solved easily by machine, and the results of this computation on the Cornell Computing Center CDC 1604 are given in table II below. We note here that the behavior of the maximizing values of ρ_q and β_q (say) as $q \to \infty$ are easily discernible from (4.1). A routine analysis shows that $\beta_q = hq^{-2} + o(q^{-2})$ and $\rho_q = \rho^* + o(1)$ where $0 < q < \infty$ and $0 < \rho^* < 1$ and where h and ρ^* maximize the coefficient of q in (4.1):

$$\text{(4.2)} \qquad \begin{aligned} h^{-1} + (1 - \rho^{*6})/6 &= 0, \\ 2/\rho^* - 4\rho^*/(1 - \rho^{*2}) + h\rho^{*5} &= 0. \end{aligned}$$

Thus, as $q \to \infty$,

$$\text{(4.3)} \qquad \begin{aligned} \rho_q^2 \sim \rho^{*2} &= (3^{1/2} - 1)/2 = .3660254, \\ \beta_q \sim hq^{-2} &= 4q^{-2}(1 + 3^{-1/2}) = 6.309401q^{-2}. \end{aligned}$$

A finer analysis can be used to produce further terms in an asymptotic expansion.

We digress in this paragraph to derive a result which was used in section 3.2. The matrix $M(\xi)$ may be inverted explicitly for ξ of the form we have been considering, the answer being expressed in terms of ρ_q and β_q. This allows one to write an expression for

$$\text{(4.4)} \qquad d^*(r, \xi^*) = \frac{(1 - \beta_q)(1 - r^2)^2 + \beta_q(r^2 - \rho_q^2)^2}{(1 - \beta_q)\beta_q(1 - \rho_q^2)^2} + q\,\frac{(1 - \beta_q)r^2(1 - r^2)^2 + \beta_q r^2\rho_q^2(r^2 - \rho_q^2)^2}{(1 - \beta_q)\beta_q\rho_q^2(1 - r_q^2)^2} + \frac{(q + 2)(q - 1)}{2}\,\frac{r^4}{(1 - \beta_q) + \beta_q\rho_q^4} + \frac{(q + 4)q(q - 1)}{6}\,\frac{r^6}{(1 - \beta_q) + \beta_q\rho_q^6}.$$

Taking $r = 1$ gives (since $d^*(1, \xi^*) = k$)

$$\text{(4.5)} \qquad \frac{(q + 1)(q + 2)(q + 3)}{6} = \frac{q + 1}{1 - \beta_q} + \frac{(q + 2)(q - 1)}{2(1 - \beta_q) + \beta_q\rho_q^4} + \frac{(q + 4)q(q - 1)}{6(1 - \beta_q) + \beta_q\rho_q^6}.$$

Equation (4.5) was used as (3.16) in section 3.2 to show that optimum designs for cubic regression on the q-cube could not be rotatable.

The following table of numbers for β_q and ρ_q were computed as described above. (Of course, for $q = 1$ we have the Guest-Hoel design.)

TABLE II

q	β_q	$q^2\beta_q$	ρ_q^2
1	0.5000	0.500	0.2000
2	0.3077	1.231	0.2657
3	0.2455	2.210	0.2970
4	0.1695	2.712	0.3142
5	0.1241	3.102	0.3249
6	0.09483	3.414	0.3321
7	0.07490	3.670	0.3373
8	0.06068	3.884	0.3412
9	0.05019	4.065	0.3442
10	0.04221	4.221	0.3465
10^2	0.6020×10^{-3}	6.020	0.364381
10^3	0.6279×10^{-5}	6.279	0.365866
10^4	0.6306×10^{-7}	6.306	0.366010
10^5	0.6309×10^{-9}	6.309	0.366024
∞		6.309401	0.3660254

From a practical point of view, what is important for other examples (for instance, larger m on the ball) is the indication that the use of the limiting values h and ρ^* for fairly small values of q leads to a value of $\max_x d(x, \xi)$ which is not too large; this aspect deserves further machine study in other contexts.

For $m \geq 4$, the same approach can be used, but of course the larger number of parameters makes the analysis messier, especially if $q > 2$. We remark that when $m = 4$, $q = 2$, the optimum weights are $\frac{1}{15}$, 0.343912, and 0.589422, at $r^2 = 0$, 0.460249, and 1, respectively.

In order to construct implementable optimum designs on the unit ball for any m, we replace the uniform distribution on each spherical shell by a distribution on a finite subset of the same shell, with the same moments. It will suffice to consider the shell of radius one. If the measure γ assigns mass α to each of the 2^q points having coordinates $\pm q^{-1/2}$, and mass β to each of the $2q$ points $(\pm 1, 0, \cdots, 0)$, $(0, \pm 1, \cdots, 0)$, $\cdots$, $(0, 0, \cdots, \pm 1)$, then the values $\alpha = q(q+2)^{-1}2^{-q}$, $\beta = 1/q(q+2)$ satisfy

$$\alpha + \beta = 1,$$

$$\int x_1^2\gamma(dx) = 1/q, \tag{4.6}$$

$$\int x_1^2x_2^2\gamma(dx) = 1/q(q+2),$$

$$\int x_1^4\gamma(dx) = 3/q(q+2).$$

Clearly all other moments of order >0 and <4 are zero.

The set of 2^q points with all coordinates $\pm q^{-1/2}$ may be replaced by a subset which is an orthogonal array of strength 4 consisting of $0(q^3)$ points, by (3.43). Thus we know how to construct optimum designs on $0(q^3)$ points. As in section 3.4, if the order of (3.42) were attainable, we could achieve $0(q^2)$ here; and perhaps less symmetric points and weights can also help to achieve a lower order than $0(q^3)$.

Using this method of construction with orthogonal arrays, and data provided by E. Seiden and described in section 3.4, we obtain the following table III giving the number of points of support for ξ in optimum designs for $q = 3, \cdots, 17$, when $m = 2$. There being one point at the origin, we obtain a design on

$$2q + 1 + n_q \tag{4.7}$$

points in this case. The values, of course, compare favorably with those of table I, where the same values of k and p apply.

TABLE III

q	(4.7)
3	23
4	25
5	43
6	45
7	79
8	81
9	147
10	149
11	151
12	281
13	283
14	285
15	287
16	289
17	291

That these designs may not be the best possible, even among designs of quite symmetric construction, is illustrated by an example of Box and Behnken [2]. Using their construction for $q = 7$, one obtains a design with 56 points on the unit sphere plus one at the origin, for a total of 57 points.

REFERENCES

[1] R. C. Bose and K. A. Bush, "Orthogonal arrays of strength two and three," *Ann. Math. Statist.*, Vol. 23 (1952), pp. 508–524.

[2] G. E. P. Box and D. W. Behnken, "Simplex-sum designs: a class of second order rotatable designs derivable from those of first order," *Ann. Math. Statist.*, Vol. 31 (1960), pp. 838–864.

[3] P. G. Guest, "The spacing of observations in polynomial regression," *Ann. Math. Statist.*, Vol. 29 (1958), pp. 294–299.

[4] P. G. Hoel, "Efficiency problems in polynomial regression," *Ann. Math. Statist.*, Vol. 29 (1958), pp. 1134–1145.

[5] J. Kiefer, "Optimum experimental designs," *J. Roy. Soc. Ser. B*, Vol. 21 (1959), pp. 273–319.

[6] ———, "Optimum experimental designs V, with applications to systematic and rotatable designs," *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 1961, Vol. 1, pp. 381–405.

[7] ———, "Optimum designs in regression problems II," *Ann. Math. Statist.*, Vol. 32 (1961), pp. 298–325.

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

[9] K. Kono, "Optimum design for quadratic regression on the k-cube," *Mem. Fac. Sci. Kyushu Univ. Ser. A.*, Vol. 16 (1962), pp. 114–122.

[10] E. Seiden and R. Zemach, "On orthogonal arrays," to be published in 1965.

[11] H. Uranisi, written communication concerning the special cubic on the q-simplex, 1962.

[12] R. Zemach, "On orthogonal arrays of strength four and their application," Michigan State University N. I. H. Report.