The Annals of Statistics
1987, Vol. 15, No. 3, 942–957

FURTHER CHARACTERIZATIONS OF DESIGN OPTIMALITY AND ADMISSIBILITY FOR PARTIAL PARAMETER ESTIMATION IN LINEAR REGRESSION

BY NORBERT GAFFKE

University of Augsburg

The paper gives a contribution to the problem of finding optimal linear regression designs, when only s out of k regression parameters are to be estimated. Also, a treatment of design admissibility for the parameters of interest is included. Previous results of Kiefer and Wolfowitz (1959), Karlin and Studden (1966) and Atwood (1969) are generalized. In particular, a connection to Tchebycheff-type approximation of $\mathbb{R}^s$-valued functions is found, which has been known in case $s = 1$. Strengthened versions of the results are obtained for invariant designs in situations, when invariance properties of the regression setup are available. Applications are given to multiple quadratic regression and to one-dimensional polynomial regression.

1. Introduction. Consider a linear regression $y(x) = a'f(x)$, $x \in \mathscr{X}$, with an unknown coefficient vector $a \in \mathbb{R}^k$, and a known $\mathbb{R}^k$-valued function f on the experimental region $\mathscr{X}$. The range $f(\mathscr{X})$ is assumed to be compact. As usual, observations $Y_1, \dots, Y_n$ on the dependent variable y, taken at points $x_1, \dots, x_n \in \mathscr{X}$, respectively, are assumed to have expectations $y(x_i)$, $1 \le i \le n$, common variance $\sigma^2 > 0$ and to be pairwise uncorrelated. Suppose that the experiment aims at estimating Ka, where K is a given $(s \times k)$-matrix of rank s. For example, when the first s components of a are to be estimated, one has $K = [I_s, 0]$, where I_s denotes the unit matrix of order s. We will study design optimality and admissibility for estimating Ka, within the approximate design theory. An approximate design ξ is a probability measure on $\mathscr{X}$ with finite support. The information matrix (per observation and unit of variance) of ξ is

$$M(\xi) = \int_{\mathscr{X}} f(x)f(x)'\,d\xi(x) = \sum_{x \in \operatorname{supp}(\xi)} f(x)f(x)'\xi(x),$$

where $\operatorname{supp}(\xi)$ denotes the support of ξ. The reduced information matrix of ξ for Ka will be denoted by $J(\xi)$, which has usually been considered only for those ξ under which Ka is estimable, i.e., $\operatorname{range}(K') \subset \operatorname{range}(M(\xi))$. In this case

$$J(\xi) = (KM^-(\xi)K')^{-1}, \tag{1}$$

where $M^-(\xi)$ is any g-inverse of $M(\xi)$. We will use another representation of $J(\xi)$, which also provides an extension of $J(\xi)$ to the set of all designs ξ,

$$J(\xi) = \min_L LM(\xi)L', \tag{2}$$

Received August 1985; revised November 1986.
AMS 1980 *subject classifications*. Primary 62K05; secondary 49B40.
Key words and phrases. Gauss–Markov estimator, approximate design theory, optimality criterion, equivalence theorem, minimax theorem, Tchebycheff approximation, invariant design, multiple quadratic regression, polynomial regression.

where the minimum in (2) is taken over all left inverses L of K' (i.e., $LK' = I_s$) and refers to the Löwner semiordering of symmetric matrices ($C \le D$ iff $D - C$ is nonnegative definite). The minimum in (2) was shown to exist by Krafft (1983). If Ka is estimable under ξ, then (1) and (2) coincide, which was shown by Sibson (1974), page 682, for $K = [I_s, 0]$, and by Gaffke and Krafft (1982), Theorem 2.4, for general K. Also, for any ξ, (2) coincides with the closure of the matrix function (1), as considered by Müller-Funk, Pukelsheim and Witting (1985), page 23. As far as design optimality is concerned, only those designs will be considered, under which Ka is estimable, or, equivalently, for which $J(\xi)$ is nonsingular. An optimality criterion is a real function ϕ on the set $PD(s)$ of all positive definite $(s \times s)$-matrices. Throughout we assume, that ϕ is convex and decreasing, where the latter means that $C, D \in PD(s)$, $C \le D$, imply $\phi(C) \ge \phi(D)$. Well-known examples are Kiefer's ϕ_p-criteria, $-1 \le p \le \infty$,

$$\phi_p(C) = \left\{s^{-1}\mathrm{tr}(C^{-p})\right\}^{1/p}, \qquad C \in PD(s)$$

[cf. Kiefer (1974), (4.18) and Kiefer (1975), page 337]. The class of information functionals considered by Pukelsheim (1980) and Pukelsheim and Titterington (1983) is included by taking $\phi = -\log j$ on $PD(s)$ (or $\phi = 1/j$, or simply $\phi = -j$), when j is an information functional. A design ξ_0 is called ϕ-optimal for Ka, iff Ka is estimable under ξ_0 and ξ_0 minimizes $\phi(J(\xi))$ over the set of all ξ, under which Ka is estimable. A new equivalence theorem to this minimization problem is presented in Section 2, and a strengthened version for invariant designs is proved. Design admissibility for Ka will be studied in Section 3. In analogy to the full parameter case, an admissible design ξ_0 for Ka is one, whose reduced information matrix $J(\xi_0)$ is a maximal element in the set of all $J(\xi)$, w.r.t. the Löwner semiordering. A necessary condition and a sufficient condition for a design to be admissible for Ka are established, and special attention is given to invariant designs. As applications of the results, three examples are treated in Section 4. Two of them deal with multiple quadratic regression on the cube $[-1, 1]^q$, where either the linear or the quadratic coefficients are to be estimated. The other is devoted to a one-dimensional polynomial regression of degree $m \ge 2$ on $[-1, 1]$ and to the estimation of the two highest coefficients. ϕ_p-optimal designs and admissible invariant designs are constructed.

2. Design optimality. Let ϕ be a given convex and decreasing optimality criterion. The set of all designs ξ, under which Ka is estimable, is assumed to be nonempty. By (2), for any ξ from this set,

$$\phi(J(\xi)) = \max_L \phi(LM(\xi)L').$$

So, the problem of minimizing $\phi(J(\xi))$ may be viewed as a minimax problem. It was shown in Gaffke (1985b), Theorem 3.2, that if there exists a ϕ-optimal design for Ka, then

$$\min_\xi \max_L \phi(LM(\xi)L') = \max_L \inf_\xi \phi(LM(\xi)L'), \tag{3}$$

where ξ ranges over the set of *all* designs, L ranges over the set of all left inverses of K' and $\phi(C)$ is defined to be ∞, if C is nonnegative definite and singular. Note, that for any fixed L the infimum on the right-hand side of (3) would be the ϕ-optimum value for designing a fictitious regression $y_L(x) = \alpha' f_L(x)$, $x \in \mathscr{X}$, where $f_L = Lf$, and the full parameter vector $\alpha \in \mathbb{R}^s$ is to be estimated. A maximin solution L_0 of (3) makes this value as bad as possible, so L_0 may be called a "least favourable" left inverse of K'. The result we will state next is, in essence, the interplay between a minimax solution ξ_0 and a maximin solution L_0 of (3). The notion of subgradients will be used, as in Pukelsheim and Titterington (1983) and Gaffke (1985a, 1985b). A symmetric $(s \times s)$-matrix E is called a subgradient of ϕ at $C_0 \in PD(s)$, iff

$$\phi(C) - \phi(C_0) \geq \operatorname{tr}\{E(C - C_0)\} \quad \text{for all } C \in PD(s).$$

By convexity of ϕ, there exists a subgradient E of ϕ at $C_0 \in PD(s)$, and E is unique iff ϕ is differentiable at C_0, in which case E is the gradient of ϕ at C_0. Since ϕ is decreasing, we have that $-E$ is nonnegative definite for any subgradient E. If C_0 is nonnegative definite and singular, then no subgradient of ϕ at C_0 exists, according to the definition $\phi(C_0) = \infty$. For $\phi = \phi_p$, $-1 \leq p < \infty$, the gradient at $C_0 \in PD(s)$ is a negative scalar multiple of C_0^{-p-1}. For $\phi = \phi_\infty$ the set of subgradients at $C_0 \in PD(s)$ is a negative scalar multiple of the convex hull of the set of all zz', where z is any normalized eigenvector for the minimum eigenvalue of C_0 [cf. Gaffke (1985a), Example 2].

THEOREM 1. *For any design ξ_0, the following conditions* (i), (ii) *and* (iii) *are equivalent*:

(i) *ξ_0 is ϕ-optimal for Ka.*

(ii) *There exists a left inverse L_0 of K', such that*

$$\phi(L_0 M(\xi_0) L_0') = \min_{\xi} \phi(L_0 M(\xi) L_0'), \tag{4}$$

and the components of $L_0 f$ are orthogonal to $\{b'f\colon\ b \in \text{nullspace}(K)\}$ in $\mathscr{L}^2(\xi_0)$-space, i.e.,

$$\int_{\mathscr{X}} L_0 f(x)(b'f(x))\, d\xi_0(x) = 0, \quad \textit{for all } b \in \text{nullspace}(K). \tag{5}$$

(iii) *There exist a left inverse L_0 of K' and a subgradient E of ϕ at $C_0 = L_0 M(\xi_0) L_0'$, such that*

(4a) *each support point of ξ_0 maximizes $(L_0 f(x))'(-E) L_0 f(x)$, $\quad x \in \mathscr{X}$,*

and (5) *from above holds true.*

PROOF. Assume (i). As in the proof of Theorem 3.2 in Gaffke (1985b) one concludes, that there exists a g-inverse $M^-(\xi_0)$, such that, taking $L_0 = J(\xi_0) K M^-(\xi_0)$, condition (4) holds true. For this L_0, we have

$$L_0 M(\xi_0) b = J(\xi_0) K b = 0, \quad \text{for all } b \in \text{nullspace}(K),$$

and hence (5).

Conversely, assume (ii). Since we have assumed that there exists a design ξ with nonsingular $J(\xi)$, and hence, by (2), with nonsingular $L_0 M(\xi) L_0'$, we have by (4) that $L_0M(\xi_0)L_0'$ is nonsingular. It can easily be seen, that nullspace(K) = range$(I_k - L_0'K)$, and hence, by (5), $L_0M(\xi_0)(I_k - L_0'K) = 0$ and $K = (L_0M(\xi_0)L_0')^{-1}L_0M(\xi_0)$. In particular, range$(K') \subset$ range$(M(\xi_0))$ and $J(\xi_0) = (KM(\xi_0)^- K')^{-1} = L_0M(\xi_0)L_0'$. If ξ is any design with range$(K') \subset$ range$(M(\xi))$, then, by (2) and (4),

$$\phi(J(\xi)) \geq \phi(L_0M(\xi)L_0') \geq \phi(L_0M(\xi_0)L_0') = \phi(J(\xi_0)),$$

and hence (i).

Thus, we have proved the equivalence of (i) and (ii). Now, the existence of a subgradient E of ϕ at $C_0 = L_0M(\xi_0)L_0'$ satisfying (4a) is equivalent to (4) [see, e.g., Gaffke (1985b), Theorem 3.1, or, for a differentiable ϕ, see Kiefer (1974), Theorem 1]. So (ii) and (iii) are equivalent. □

REMARK 1. The equivalence of (i) and (ii) of Theorem 1 strengthens and generalizes Theorem 3.4 of Atwood (1969), who gave a sufficient condition for D-optimality for Ka in the case $K = [I_s, 0]$ (D_s-optimality). A special case of his result was rediscovered by Näther and Reinsch (1981), Theorem 5. For $K = [I_s, 0]$ a left inverse L_0 of K' has the form $L_0 = [I_s, B_0]$, with some $s \times (k-s)$-matrix B_0. Hence $L_0 f = f^{(1)} + B_0 f^{(2)}$, where $f^{(1)}$ and $f^{(2)}$ consist of the first s and the last $k-s$ components of f, respectively, and the space $\{b'f\colon\ b \in \text{nullspace}(K)\}$ is generated by the components of $f^{(2)}$. In Atwood's theorem our condition (5) is replaced by the stronger condition that

$$(5') \qquad f^{(2)}(x) = 0, \quad \text{for all } x \in \text{supp}(\xi_0),$$

and (4), with the D-criterion ϕ_0 and (5′) together were proved to be sufficient for D_s-optimality of ξ_0. Actually, Atwood gave a slightly more general version of this, employing a matrix $L_0 = [A_0, B_0]$, where A_0 is a nonsingular $(s \times s)$-matrix, but for the D-criterion this is a straightforward extension.

REMARK 2. The equivalence of (i) and (iii) of Theorem 1 gives an analogue for $s \geq 1$ to the result of Kiefer and Wolfowitz (1959) for $s = 1$, [Theorem 1 of their paper; see also, Kiefer (1959), page 301]. For $s = 1$ and $K = c'$, where $c \in \mathbb{R}^k$, $c \neq 0$, there is essentially one optimality criterion, $\phi(\gamma) = 1/\gamma$, $\gamma > 0$, and ϕ-optimality of a design for $c'a$ is simply called c-optimality. In this case, the subgradient E occurring in (iii) of our Theorem 1 is merely a negative real number, which can be ignored in (4a), and one gets back the Kiefer–Wolfowitz result on c-optimality. Moreover, Kiefer and Wolfowitz (1959) pointed out, that for $s = 1$ the $(s \times k)$ = matrix L_0 occurring in (iii) is a Tchebycheff coefficient vector. This has an analogue for $s \geq 1$. For, if L_0 and E are as in (iii), then for any left inverse L of K' we have

$$\begin{aligned}\max_x (L_0 f(x))'(-E)L_0 f(x) &= \int_{\mathscr{X}} (L_0 f(x))'(-E)L_0 f(x)\, d\xi_0(x)\\ &\leq \int_{\mathscr{X}} (Lf(x))'(-E)Lf(x)\, d\xi_0(x)\\ &\leq \max_x (Lf(x))'(-E)Lf(x).\end{aligned}$$

Hence L_0 is a Tchebycheff coefficient matrix in the sense that

$$(6) \qquad \max_x (L_0 f(x))'(-E) L_0 f(x) = \min_L \max_x (Lf(x))'(-E) Lf(x),$$

where the maximum in (6) is taken over all $x \in \mathscr{X}$, and the minimum on the right-hand side of (6) is taken over all left inverses L of K'. Writing each left inverse $L = (K')^- - D$, where $(K')^-$ is any fixed left inverse of K', and D is an $(s \times k)$-matrix, depending on L, with $DK' = 0$, (6) becomes a Tchebycheff type approximation problem: Find a best approximation of the $\mathbb{R}^s$-valued function $(K')^- f$ by an element from the space $\{Df\colon DK' = 0\}$ with respect to the Tchebycheff type seminorm

$$\|h\| = \max_x \{h(x)'(-E)h(x)\}^{1/2},$$

where h is an $\mathbb{R}^s$-valued function on $\mathscr{X}$ with compact range. However, in case $s > 1$ the approximation problem (6) depends on the subgradient E from (iii), which is not known in advance. This prevents a direct use of (6) for finding the correct left inverse L_0 for (iii) and constructing an optimal design ξ_0 from (4a) and (5). This is in contrast to the case $s = 1$. A study of Tchebycheff-type approximation problems as in (6) will be done elsewhere. A related problem has been considered by Studden (1980b).

REMARK 3. The equivalence of (i) and (iii) of Theorem 1 is closely related to Corollary 7.2 of Pukelsheim (1980). In fact, our result can be derived from Corollary 7.2 of his paper, when the optimality criterion is given by an information functional, but it does require some steps.

Now we will consider the case, that there are invariance (or equivariance) properties of the regression setup $y(x) = a'f(x)$, of the parametric function Ka, and of the optimality criterion ϕ. Let G be a group of one-to-one transformations g from the experimental region on itself, let $\mathscr{Q}$ be a compact group of nonsingular $(k \times k)$-matrixes Q, and let there be given a surjective mapping Q_g, $g \in G$, from G on $\mathscr{Q}$. As in Pukelsheim (1987), Section 3.1, we assume the following (α), (β) and (γ):

(α) $f(g(x)) = Q_g f(x)$ for all $x \in \mathscr{X}$, $g \in G$.
(β) range(QK') = range(K'), or, equivalently, $QK'\overline{Q}^{-1} = K'$, where

$$\overline{Q} = (KK')^{-1}KQK', \quad \text{for all } Q \in \mathscr{Q}.$$

(γ) $\phi(\overline{Q}C(\overline{Q})') = \phi(C)$ for all $Q \in \mathscr{Q}$, $C \in PD(s)$.

A design ξ will be called $\mathscr{Q}$-invariant, iff its information matrix $M(\xi)$ is invariant under the group of congruence transformations given by $\mathscr{Q}$, i.e., $QM(\xi)Q' = M(\xi)$ for all $Q \in \mathscr{Q}$. Of course, $\mathscr{Q}$-invariance is a weaker property than G-invariance, which would mean that $\xi^g = \xi$ for all $g \in G$, where ξ^g denotes the image of ξ under g [ξ^g is supported by the image of supp(ξ) under g, with weights $\xi^g(g(x)) = \xi(x)$]. If G is a finite group, as in most applications, then a design ξ is $\mathscr{Q}$-invariant iff there exists a G-invariant design $\bar{\xi}$, such that

$M(\xi) = M(\bar{\xi})$. For finding a ϕ-optimal design for Ka, we may restrict to $\mathscr{Q}$-invariant designs. In fact, if ξ is any design, under which Ka is estimable, then by taking the average of $QM(\xi)Q'$ over $Q \in \mathscr{Q}$ with respect to the Haar probability measure on $\mathscr{Q}$, we get an $\overline{M}$, which is an element of the convex hull of the orbit $\{QM(\xi)Q': Q \in \mathscr{Q}\}$. This orbit is contained in the set of all information matrices, and so is its convex hull. Hence there exists a design ξ_1 with $M(\xi_1) = \overline{M}$, and, clearly, ξ_1 is $\mathscr{Q}$-invariant. As in Pukelsheim (1987), Section 3.2, one concludes that $\phi(J(\xi_1)) \le \phi(J(\xi))$. So, if there exists a ϕ-optimal design for Ka, then there exists a $\mathscr{Q}$-invariant design, which is ϕ-optimal for Ka. For finding a $\mathscr{Q}$-invariant, ϕ-optimal design for Ka, the following complement to Theorem 1 will be useful, which reduces the variability of the unknown quantities L_0 and E of Theorem 1, in that it is shown that these can be chosen to be invariant.

Theorem 1(a). *Under assumptions* (α), (β) *and* (γ), *let* ξ_0 *be a* $\mathscr{Q}$-*invariant*, ϕ-*optimal design for* Ka. *Then there exist a left inverse* L_0 *of* K' *and a subgradient* E *of* ϕ *at* $C_0 = L_0 M(\xi_0) L_0'$, *which are invariant in the sense that*

$$\overline{Q}^{-1} L_0 Q = L_0 \quad \text{and} \quad (\overline{Q})' E \overline{Q} = E, \quad \text{for all } Q \in \mathscr{Q},$$

and which satisfy (4), (4a) *and* (5) *of Theorem* 1.

Proof. Consider the function $\Phi(M) = \phi((KM^- K')^{-1})$, which is defined on the convex cone of all nonnegative definite $(k \times k)$-matrices M with range$(K') \subset$ range(M). By Theorem 3.1 and Lemma 3.2 of Gaffke (1985b), there exists a subgradient B of Φ at $M(\xi_0)$, such that

$$f(x)'(-B)f(x) \le \operatorname{tr}\{-BM(\xi_0)\}, \quad \text{for all } x \in \mathscr{X}. \tag{7}$$

For any $Q \in \mathscr{Q}$, the matrix $B(Q) = Q'BQ$ again is a subgradient of Φ at $M(\xi_0)$ and satisfies (7), and so does the average B_0 of all $B(Q)$ (w.r.t. the Haar probability measure on $\mathscr{Q}$). By Lemma 3.2 of Gaffke (1985b), there exist a g-inverse $M^-(\xi_0)$, a subgradient E of ϕ at $J(\xi_0)$ and a nonnegative definite $(k \times k)$-matrix D with $DM(\xi_0) = 0$, such that

$$B_0 = (M^-(\xi_0))'K'J(\xi_0)EJ(\xi_0)KM^-(\xi_0) - D.$$

The subgradient $E = KB_0K'$ is invariant, i.e., $(\overline{Q})'E\overline{Q} = E$ for all $Q \in \mathscr{Q}$. Take $L = J(\xi_0)KM^-(\xi_0)$. We have thus found an invariant subgradient E of ϕ at $J(\xi_0)$, and a left inverse L of K', such that

$$(Lf(x))'(-E)Lf(x) \le \operatorname{tr}\{-EJ(\xi_0)\}, \quad \text{for all } x \in \mathscr{X}, \tag{8}$$

and

$$LM(\xi_0)b = 0, \quad \text{for all } b \in \text{nullspace}(K). \tag{9}$$

Now, (8) and (9) hold true with L replaced by $L(Q) = \overline{Q}^{-1}LQ$, for any $Q \in \mathscr{Q}$, and hence for the average L_0 of all $L(Q)$. From (9) one gets $J(\xi_0) = L_0 M(\xi_0) L_0'$. So, L_0 and E are the invariant quantities satisfying (4a) and (5) of Theorem 1. Equation (4) is a consequence of (4a). □

3. Design admissibility. Admissibility of approximate designs for linear unbiased estimation in linear regression was considered by Kiefer (1959), Section 3, Karlin and Studden (1966), Section 7, and Pukelsheim (1980), pages 359–360. A design ξ^* is said to be better than a design ξ, iff $M(\xi) \le M(\xi^*)$ and $M(\xi) \ne M(\xi^*)$. A design ξ_0 is said to be admissible, iff there does not exist a design which is better than ξ_0. Another formulation for this in terms of variances of Gauss–Markov estimators is the following. For a design ξ and a coefficient vector $c \in \mathbb{R}^k$ let $V(\xi, c) = c'M^-(\xi)c$, if $c \in \text{range}(M(\xi))$, and $V(\xi, c) = \infty$, otherwise. If ξ corresponds to an exact n point design, then $(\sigma^2/n)V(\xi, c)$ is the variance of the Gauss–Markov estimator for $c'a$ under ξ, which we define to be infinity, if $c'a$ is not estimable under ξ. By Lemma 2 of Stepniak, Wang and Wu (1984), a design ξ^* is better than a design ξ, iff $V(\xi^*, c) \le V(\xi, c)$ for all $c \in \mathbb{R}^k$, with strict inequality for at least one $c_0 \in \mathbb{R}^k$. For partial parameter estimation, when the parameters of interest are given by the $\mathbb{R}^s$-valued function Ka, we will restrict to linear parametric functions $t'Ka$, where $t \in \mathbb{R}^s$. This suggests the following definition.

Definition. A design ξ^* is said to be better than a design ξ for Ka, iff $V(\xi^*, K't) \le V(\xi, K't)$ for all $t \in \mathbb{R}^s$, with strict inequality for at least one $t_0 \in \mathbb{R}^s$. A design ξ_0 is said to be admissible for Ka, iff there does not exist a design which is better than ξ_0 for Ka.

In analogy to the full parameter case mentioned above, a reformulation can be given in terms of the reduced information matrices from (2). The proof of the following lemma is omitted.

Lemma. *A design ξ^* is better than a design ξ for Ka, if and only if $J(\xi) \le J(\xi^*)$ and $J(\xi) \ne J(\xi^*)$.*

The following result provides an extension of Theorem 7.1 of Karlin and Studden (1966) to the partial parameter case.

Theorem 2. *Let ξ_0 be a given design. Then:*

(a) *If ξ_0 is admissible for Ka, then there exist a nonnegative definite, nonzero $(s \times s)$-matrix A and a left inverse L_0 of K', such that*

(10) *each support point of ξ_0 maximizes* $(L_0 f(x))' A L_0 f(x), \quad x \in \mathscr{X},$

and

$$\int_{\mathscr{X}} L_0 f(x)(b'f(x))\, d\xi_0(x) = 0, \quad \text{for all } b \in \text{nullspace}(K). \tag{11}$$

(b) *If there exist a positive definite $(s \times s)$-matrix A and a left inverse L_0 of K', such that* (10) *and* (11) *from above hold true, then ξ_0 is admissible for Ka.*

PROOF. (a) Consider the set of all nonnegative definite $(s \times s)$-matrices B with $\operatorname{tr} B = 1$. By (2), for any B from this set, and any design ξ,

$$\operatorname{tr}\{B(J(\xi) - J(\xi_0))\} = \min_L \operatorname{tr}\{B(LM(\xi)L' - J(\xi_0))\},$$

and this is a concave and upper semicontinuous function of $M(\xi)$. So we can conclude as Karlin and Studden (1966), page 808. Admissibility of ξ_0 for Ka implies

$$0 \geq \max_\xi \min_B \operatorname{tr}\{B(J(\xi) - J(\xi_0))\}, \tag{12}$$

and by a well-known minimax result [see, e.g., Parthasarathy and Raghavan (1971), Theorem 5.3.6], the maximum and the minimum on the right-hand side of (12) can be interchanged. Hence there exists a nonnegative definite $(s \times s)$-matrix A with $\operatorname{tr} A = 1$, such that

$$\operatorname{tr}\{AJ(\xi_0)\} = \max_\xi \operatorname{tr}\{AJ(\xi)\}.$$

Again, by (2) and by the general minimax result,

$$\min_L \operatorname{tr}\{ALM(\xi_0)L'\} = \max_\xi \min_L \operatorname{tr}\{ALM(\xi)L'\}$$
$$= \inf_L \max_\xi \operatorname{tr}\{ALM(\xi)L'\}.$$

The last infimum is attained at some left inverse L_1 of K', since the function

$$\eta(L) = \max_\xi [\operatorname{tr}\{ALM(\xi)L'\}]^{1/2}$$

is a seminorm on the space of all $(s \times k)$-matrices L, and the set of all left inverses of K' is an affine subset of this space. So, the two-person zero-sum game with payoff for player 1

$$\kappa(\xi, L) = \operatorname{tr}\{ALM(\xi)L'\},$$

where player 1 chooses a design ξ and player 2 chooses a left inverse L of K', is definite, ξ_0 is a maximin strategy for player 1, and there exists a minimax strategy L_1 for player 2. Hence

$$\operatorname{tr}\{AL_1M(\xi_0)L_1'\} = \max_\xi \operatorname{tr}\{AL_1M(\xi)L_1'\}, \tag{13}$$

and

$$\operatorname{tr}\{AL_1M(\xi_0)L_1'\} = \min_L \operatorname{tr}\{ALM(\xi_0)L'\}. \tag{14}$$

It can easily be seen, that (14) implies $AL_1M(\xi_0)P = 0$, where P denotes the orthogonal projector from $\mathbb{R}^k$ on nullspace(K). Consider the system of matrix equations

$$AX = AL_1, \qquad XK' = I_s, \qquad XM(\xi_0)P = 0.$$

By Theorem 2.3.3 of Rao and Mitra (1971), this system has a solution for X, since, as it can readily be checked, range$(AL_1) \subset$ range(A) (trivially),

$$\text{nullspace}([K', M(\xi_0)P]) \subset \text{nullspace}([I_s, 0])$$

and

$$A[I_s, 0] = AL_1[K', M(\xi_0)P].$$

Taking $L_0 = X$, one gets a left inverse of K', such that, by (13),

$$\mathrm{tr}\{AL_0M(\xi_0)L_0'\} = \max_\xi \mathrm{tr}\{AL_0M(\xi)L_0'\},$$

which is the same as (10), and $L_0M(\xi_0)P = 0$, which is the same as (11).

(b) Let ξ be a design with $J(\xi_0) \le J(\xi)$. By (11) $L_0M(\xi_0)L_0' = J(\xi_0)$, and hence, by (10) and (2),

$$\begin{aligned}\mathrm{tr}\{AJ(\xi_0)\} &= \mathrm{tr}\{AL_0M(\xi_0)L_0'\}\\ &\ge \mathrm{tr}\{AL_0M(\xi)L_0'\} \ge \mathrm{tr}\{AJ(\xi)\}.\end{aligned}$$

Since A is positive definite, this implies $J(\xi_0) = J(\xi)$, and hence ξ_0 is admissible for Ka. □

REMARK 4. Our previous remark on the Tchebycheff approximation (Section 2, Remark 2), applies to the present situation as well. If ξ_0, A and L_0 are such that (10) and (11) hold true, then L_0 is a Tchebycheff coefficient matrix, in the sense that it minimizes

$$\max_x (Lf(x))'ALf(x),$$

over the set of all left inverses L of K'.

Under the invariance assumptions (α) and (β) from Section 2, the class of $\mathcal{Q}$-invariant designs, which are admissible for Ka, may be of interest. For finding this class, the following result complementing Theorem 2 will be useful.

THEOREM 2(a). *Under the assumptions* (α) *and* (β) *from Section* 2, *let* ξ_0 *be a* $\mathcal{Q}$-*invariant design, which is admissible for* Ka. *Then there exist a nonnegative definite, nonzero* $(s \times s)$-*matrix* A *and a left inverse* L_0 *of* K', *which are invariant in the sense that*

$$(\overline{Q})'A\overline{Q} = A \quad \textit{and} \quad \overline{Q}^{-1}L_0Q = L_0, \quad \textit{for all } Q \in \mathcal{Q},$$

and such that (10) *and* (11) *of Theorem* 2 *hold true*.

PROOF. First, as in the proof of Theorem 2, one concludes that there exists a nonnegative definite, nonzero $(s \times s)$-matrix A, such that

$$\mathrm{tr}\{AJ(\xi_0)\} \ge \mathrm{tr}\{AJ(\xi)\}, \quad \text{for all } \xi. \tag{15}$$

Formula (15) holds true with A replaced by $A(Q) = (\overline{Q})'A\overline{Q}$, for any $Q \in \mathcal{Q}$, and hence for the average of all $A(Q)$ (w.r.t. the Haar probability measure on $\mathcal{Q}$). So we can assume A from (15) to be invariant. Now, proceeding further as in the proof of Theorem 2, one gets a left inverse L_0 of K', such that

$$\mathrm{tr}\{AL_0M(\xi_0)L_0'\} \ge \mathrm{tr}\{AL_0M(\xi)L_0'\}, \quad \text{for all } \xi, \tag{16}$$

and

(17) $$L_0 M(\xi_0) b = 0, \quad \text{for all } b \in \text{nullspace}(K).$$

Again, (16) and (17) hold true with L_0 replaced by $L_0(Q) = \bar{Q}^{-1} L_0 Q$, for any $Q \in \mathscr{Q}$, and hence for the average of all $L_0(Q)$ [note that (17) implies $L_0 M(\xi_0) L_0' = J(\xi_0)$, and hence the left-hand side of (16) is not affected, when replacing L_0 by $L_0(Q)$ or by their average]. In this way, one obtains an invariant left inverse of K', which, together with the invariant A, satisfies (10) and (11) of Theorem 2. □

4. Examples.

EXAMPLE 1. Consider multiple quadratic regression on the cube $\mathscr{X} = [-1, 1]^q$,

(18) $$y(x) = a' f(x) = a_0 + \sum_{i=1}^{q} a_i x_i + \sum_{1 \le i \le j \le q} a_{ij} x_i x_j,$$

$x = (x_1, \ldots, x_q)$. Suppose that the constant term a_0 and the first-order coefficients $a_1, \ldots, a_q$ are to be estimated, i.e., $K = [I_{q+1}, 0]$. Let G be the group of all permutations and the sign changes of coordinates, $G = \{g = (\pi, \varepsilon)\}$, where π is a permutation of the indices $1, \ldots, q$, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_q)$, with $\varepsilon_i \in \{-1, +1\}$, and $(\pi, \varepsilon)(x) = (\varepsilon_1 x_{\pi(1)}, \ldots, \varepsilon_q x_{\pi(q)})$. Conditions ($\alpha$) and ($\beta$) from Section 2 are satisfied, where $\mathscr{Q} = \{Q_{\pi, \varepsilon}\}$,

$$Q_{\pi, \varepsilon} = \begin{bmatrix} R_{\pi, \varepsilon} & 0 \\ 0 & S_{\pi, \varepsilon} \end{bmatrix}, \qquad \bar{Q}_{\pi, \varepsilon} = R_{\pi, \varepsilon},$$

and the quadratic matrices $R_{\pi, \varepsilon}$ and $S_{\pi, \varepsilon}$ are of order $q + 1$ and $q(q + 1)/2$, respectively, which are defined by

$$R_{\pi, \varepsilon}(t_0, t_1, \ldots, t_q)' = (t_0, \varepsilon_1 t_{\pi(1)}, \ldots, \varepsilon_q t_{\pi(q)})',$$
$$S_{\pi, \varepsilon}(t_{ij}\colon 1 \le i \le j \le q)' = (\varepsilon_i \varepsilon_j t_{\langle \pi(i), \pi(j) \rangle}\colon 1 \le i \le j \le q)'.$$

The notation $\langle \pi(i), \pi(j) \rangle$ means $(\pi(i), \pi(j))$, if $\pi(i) \le \pi(j)$, and $(\pi(j), \pi(i))$, otherwise. For convenience, we have indexed the components of $q(q + 1)/2$-dimensional vectors by pairs (i, j), $1 \le i \le j \le q$, arranged in lexicographic order, say. Below, we will employ ϕ_p-criteria, $-1 \le p \le \infty$, which satisfy condition (γ) from Section 2, since the matrices $R_{\pi, \varepsilon}$ are orthogonal. Let A be a nonnegative definite, nonzero matrix of order $q + 1$, and let L_0 be a left inverse of K', hence $L_0 = [I_{q+1}, B]$, with $B = (b_{h,(i,j)})$ having $q + 1$ rows [indexed by $h = 0, 1, \ldots, q$] and $q(q + 1)/2$ columns [indexed by (i, j), $1 \le i \le j \le q$]. The invariance conditions from Theorem 2(a) mean that A is diagonal with elements $\alpha_0 \ge 0$ and $\alpha_1 = \cdots = \alpha_q \ge 0$ in the diagonal, where $\alpha_0 + \alpha_1 > 0$, and L_0 is such that $b_{0,(i,i)} = -\beta$, say, and all the other $b_{h,(i,j)}$ are zero. For such matrices A and L_0 we have

(19) $$L_0 f(x) = (1 - \beta r^2(x), x_1, \ldots, x_q)',$$
$$(L_0 f(x))' A L_0 f(x) = \alpha_0 (1 - \beta r^2(x))^2 + \alpha_1 r^2(x),$$

where $r^2(x) = \sum_{i=1}^{q} x_i^2$.

We will determine $\mathscr{Q}$-invariant, ϕ_p-optimal designs for $(a_0, a_1, \dots, a_q)'$. These turn out to be those $\mathscr{Q}$-invariant designs, which give total weight λ_p to the set of vertices of the cube, and weight $1 - \lambda_p$ to the origin, where

$$\lambda_p = q^{1/(p+1)}/\left(1 + q^{1/(p+1)}\right), \quad \text{if } -1 < p < \infty,$$

and $\lambda_\infty = 1/2$. For $p = -1$ no ϕ_{-1}-optimal design exists, unless $q = 1$, in which case ξ_0 with support $-1, 0, 1$ and weights $1/4, 1/2, 1/4$ is ϕ_p-optimal for all $p \in [-1, \infty]$ [see also Pukelsheim (1980), page 361]. These results are obtained by Theorems 1 and 1(a) as follows. By our theorems, a $\mathscr{Q}$-invariant design ξ_0 is ϕ_p-optimal for $(a_0, a_1, \dots, a_q)'$, if and only if there exist invariant matrices A and L_0 as above, such that

$$\text{(20)} \qquad -A \text{ is a subgradient of } \phi_p \text{ at } C_0 = \int_{\mathscr{X}} L_0 f(x)(L_0 f(x))' \, d\xi_0(x),$$

$$\text{(21)} \qquad \text{each support point of } \xi_0 \text{ maximizes (19) over } x \in [-1,1]^q,$$

$$\text{(22)} \qquad \int_{\mathscr{X}} (1 - \beta r^2(x)) r^2(x) \, d\xi_0(x) = 0.$$

Equations (21) and (22) imply that each support point of ξ_0 is either the origin or a vertex of the cube, and $\beta = 1/q$, $\alpha_0 = q\alpha_1$. So, if λ is the total weight of ξ_0 assigned to the set of vertices, we have $C_0 = \operatorname{diag}(1 - \lambda, \lambda, \dots, \lambda)$. For $-1 \le p < \infty$, the gradient of ϕ_p at C_0 is a negative multiple of C_0^{-p-1}, and hence, by (20), $\{(1-\lambda)/\lambda\}^{-p-1} = q^{-1}$, and $\lambda = \lambda_p$, if $p > -1$. If $p = -1$ and $q \ge 2$, then no solution for λ exists, and hence no ϕ_{-1}-optimal design. If $p = -1$ and $q = 1$, then λ may be chosen arbitrarily from $(0, 1)$. Let $p = \infty$. The set of subgradients of ϕ_∞ at C_0 is a negative multiple of the convex hull of $\{zz'\}$, where z ranges over the set of all normalized eigenvectors for the minimum eigenvalue of C_0. Since A is nonsingular, we have from (20) that the multiplicity of the minimum eigenvalue of C_0 is $q + 1$, and hence $\lambda_\infty = 1/2$.

It can easily be checked, that the designs ξ_0, thus obtained, indeed satisfy conditions (20), (21) and (22), with $\beta = 1/q$, $\alpha_0 = q\alpha_1$ and an appropriate $\alpha_1 > 0$, and hence, by Theorem 1, they are ϕ_p-optimal.

Next, we will show that the class of all $\mathscr{Q}$-invariant, admissible designs for $(a_0, a_1, \dots, a_q)'$ consists of all ξ_0, which are $\mathscr{Q}$-invariant and $\operatorname{supp}(\xi_0) \subset S$, where S denotes the set of the vertices and the origin of the cube. If ξ_0 is $\mathscr{Q}$-invariant and admissible, then, by Theorem 2(a), there exist $\alpha_0 \ge 0$, $\alpha_1 \ge 0$, $\alpha_0 + \alpha_1 > 0$ and $\beta \in \mathbb{R}$, such that (21) and (22) hold true. These imply $\operatorname{supp}(\xi_0) \subset S$. Conversely, let ξ_0 be $\mathscr{Q}$-invariant and $\operatorname{supp}(\xi_0) \subset S$. Take $\alpha_0 = 1$, $\alpha_1 = 1/q$ and $\beta = 1/q$. Then (21) and (22) hold true, and, since $A = \operatorname{diag}(1, q^{-1}, \dots, q^{-1})$ is positive definite, ξ_0 is admissible by Theorem 2.

EXAMPLE 2. Again, we consider regression (18) on the cube $\mathscr{X} = [-1, 1]^q$, and let $q \ge 2$. Suppose now, that the second-order coefficients a_{ij}, $1 \le i \le j \le q$, are to be estimated, i.e., $K = [0, I_{q(q+1)/2}]$. With the groups $G = \{(\pi, \varepsilon)\}$ and $\mathscr{Q} = \{Q_{\pi,\varepsilon}\}$ from Example 1, conditions (α) and (β) are satisfied, where now $\overline{Q}_{\pi,\varepsilon} = S_{\pi,\varepsilon}$. Let A be a nonnegative definite, nonzero matrix of order $q(q+1)/2$, the elements of which we denote by $a_{(i,j),(k,l)}$, $1 \le i \le j \le q$, $1 \le k \le l \le q$, and let L_0 be a left inverse of K', hence $L_0 = [B, I_{q(q+1)/2}]$, with $B = (b_{(i,j),h})$, $1 \le i \le j \le q$, $0 \le h \le q$. Invariance of A and L_0, as explained in Theorems 1(a) and 2(a), means that $a_{(i,i),(i,i)} = \alpha_1$, $a_{(i,i),(j,j)} = \alpha_0$, $i \ne j$, $a_{(i,j),(i,j)} = \alpha_2$, $i < j$, and all the other elements of A are zero, and $b_{(i,i),0} = -\beta$ and all the other elements of B are zero. For such matrices A and L_0 we have $L_0 f(x) = (g_{ij}(x)\colon 1 \le i \le j \le q)'$, where $g_{ii}(x) = x_i^2 - \beta$, $g_{ij}(x) = x_i x_j$, $i < j$, and

$$(L_0 f(x))' A L_0 f(x) = \alpha_1 \sum_{i=1}^{q} \left(x_i^2 - \beta\right)^2 + \alpha_0 \sum_{i \ne j} \left(x_i^2 - \beta\right)\left(x_j^2 - \beta\right) + \alpha_2 \sum_{i<j} \left(x_i^2 x_j^2\right). \tag{23}$$

The matrix A is nonnegative definite and nonzero, which means that $\alpha_1 \ge 0$, $\alpha_2 \ge 0$, $\alpha_1 + \alpha_2 > 0$ and $-(q-1)\alpha_1 \le \alpha_0 \le \alpha_1$. By Theorems 2 and 2(a), for finding admissible $\mathcal{Q}$-invariant designs ξ_0 for Ka, we have to evaluate the following conditions:

(24) each support point of ξ_0 maximizes (23) over $x \in [-1,1]^q$,

$$\beta = q^{-1} E(r^2), \tag{25}$$

where, for short, $E(\cdot)$ means expectation w.r.t. ξ_0, and $r^2(x) = \sum_{i=1}^q x_i^2$.

It is easy to see that any x maximizing (23) is from the lattice $\{-1, 0, 1\}^q$. On this lattice, (23) is a quadratic function of $r^2 = r^2(x) \in \{0, 1, \dots, q\}$, namely,

$$\begin{aligned} h(r^2) &= \left(\frac{\alpha_2}{2} + \alpha_0\right) r^4 + \left\{\alpha_1 - 2\beta\left(\alpha_1 + (q-1)\alpha_0\right) - \left(\frac{\alpha_2}{2} + \alpha_0\right)\right\} r^2 \\ &\quad + q\beta^2\left(\alpha_1 + (q-1)\alpha_0\right) \\ &= \gamma_2 r^4 + \gamma_1 r^2 + \gamma_0, \quad \text{say.} \end{aligned}$$

Considering the maximum of $h(r^2)$, conditions (24) and (25) lead to the following:

Case 1: $\gamma_2 > 0$. Then $r^2(x) \in \{0, q\}$ on $\operatorname{supp}(\xi_0)$ and $E(r^2) \ge q/2$.

Case 2: $\gamma_2 = 0$, which implies $\gamma_1 \ge 0$. If $\gamma_1 > 0$, then $r^2(x) = q$ on $\operatorname{supp}(\xi_0)$. If $\gamma_1 = 0$, then $\operatorname{supp}(\xi_0) \subset \{-1, 0, 1\}^q$ arbitrary, but $E(r^2) \ge q/2$.

Case 3: $\gamma_2 < 0$. Then $r^2(x) \in \{t, t+1\}$ on $\operatorname{supp}(\xi_0)$, where $t \in \{1, \dots, q-1\}$, $t \ge (q-1)/2$.

So, if ξ_0 is $\mathcal{Q}$-invariant and admissible for Ka, then $\operatorname{supp}(\xi_0) \subset \{-1, 0, 1\}^q$, and $E(r^2) \ge q/2$ or $r^2(x) \in \{(q-1)/2, (q+1)/2\}$, in which case q is odd. Conversely, if ξ_0 is $\mathcal{Q}$-invariant, if $\operatorname{supp}(\xi_0) \subset \{-1, 0, 1\}^q$ and if $E(r^2) > q/2$, then ξ_0 is admissible for Ka. For, we can choose $\alpha_0, \alpha_1, \alpha_2$, such that A from above is positive definite, and such that (24) holds true with β from (25) [e.g., take $\alpha_0 = -1/(q-1)$, $\alpha_1 = 2\beta/(2\beta - 1)$ and $\alpha_2 = 2/(q-1)$].

The question remains open, whether $\mathcal{Q}$-invariant designs are admissible for Ka, which are supported by a subset of the lattice and such that $E(r^2) = q/2$ or $r^2(x) \in \{(q-1)/2, (q+1)/2\}$ on the support.

EXAMPLE 3. Consider polynomial regression of degree $m \geq 2$ on the interval $\mathcal{X} = [-1, 1]$,

$$y(x) = a'f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m. \tag{26}$$

Suppose that the two highest coefficients a_{m-1} and a_m are to be estimated, i.e., $K = [0, I_2]$. The group G consists of two elements, the identity and the reflection at zero. Obviously, conditions (α), (β) are satisfied, and for ϕ_p-criteria, $-1 \leq p \leq \infty$, (γ) is satisfied, where $\mathcal{Q} = \{Q_1, Q_2\}$, $Q_1 = I_{m+1}$, Q_2 is diagonal with elements $(-1)^i$, $i = 0, 1, \ldots, m$, and $\overline{Q}_1 = I_2$, $\overline{Q}_2 = \operatorname{diag}((-1)^{m-1}, (-1)^m)$. A design ξ is $\mathcal{Q}$-invariant, iff the odd moments of ξ up to the order $2m-1$ are zero. Let A be a nonnegative definite, nonzero (2×2)-matrix and let L_0 be a left inverse of K'. Clearly,

$$L_0 f(x) = \left(x^{m-1} - P_0(x), x^m - Q_0(x)\right)',$$

where P_0 and Q_0 are polynomials of degree at most $m-2$. Invariance of A and L_0, as explained in Theorems 1(a) and 2(a), means that A is diagonal, and one of the components of $L_0 f$ is an even polynomial and the other is odd. By our results, the problems of finding optimal or admissible $\mathcal{Q}$-invariant designs for $(a_{m-1}, a_m)'$ lead to the following. For given $\alpha_1 \geq 0$ and $\alpha_2 \geq 0$, with $\alpha_1 + \alpha_2 > 0$, find a design ξ and a pair of polynomials P_0, Q_0 each of degree at most $m-2$, such that

(27) each support point of ξ maximizes $\alpha_1\left(x^{m-1} - P_0(x)\right)^2 + \alpha_2\left(x^m - Q_0(x)\right)^2$ over $x \in [-1, 1]$,

and

$$\begin{aligned} &\int_{[-1,1]} \left(x^{m-1} - P_0(x)\right) x^i \, d\xi(x) = 0, \\ &\int_{[-1,1]} \left(x^m - Q_0(x)\right) x^i \, d\xi(x) = 0, \qquad i = 0, 1, \ldots, m-2. \end{aligned} \tag{28}$$

As pointed out in Remarks 2 and 4, any such pair (P_0, Q_0) provides a Tchebycheff approximation of the monomials x^{m-1} and x^m, namely

(29) (P_0, Q_0) minimizes

$$\max_{-1 \leq x \leq 1} \left\{\alpha_1\left(x^{m-1} - P(x)\right)^2 + \alpha_2\left(x^m - Q(x)\right)^2\right\}$$

over the set of all pairs (P, Q) of polynomials of degree at most $m-2$.

If one of α_1, α_2 is zero, then (29) is a special case of ordinary Tchebycheff approximation, the solution of which is well known.

—If $\alpha_1 = 0$, then $x^m - Q_0(x) = 2^{-(m-1)}T_m(x)$, (and P_0 arbitrary).
—If $\alpha_2 = 0$, then $x^{m-1} - P_0(x) = 2^{-(m-2)}T_{m-1}(x)$ (and Q_0 arbitrary),

where T_r denotes the rth Tchebycheff polynomial of the first kind. In case that $\alpha_1 > 0$ and $\alpha_2 > 0$ it can be shown that the unique solution of (29) is given by

$$x^{m-1} - P_0(x) = 2^{-(m-2)}T_{m-1}(x),$$

$$x^m - Q_0(x) = 2^{-(m-1)}(T_m(x) - \beta T_{m-2}(x)),$$

where $\beta = \alpha_1/\alpha_2$, if $\alpha_1 < \alpha_2$, and $\beta = 1$, otherwise. So, if $\alpha_1 = 0$ or $\alpha_2 = 0$, then (27) and (28) yield designs ξ_0 or ξ_1, which are the optimal designs for estimating a_m or a_{m-1}, respectively [see Kiefer and Wolfowitz (1959), Theorem 1]. ξ_0 is supported by the points $\cos(j\pi/m)$, $j = 0, 1, \dots, m$, with weights $1/m$, if $1 \le j \le m-1$, and $1/(2m)$, if $j = 0$ or m. ξ_1 is supported by $\cos(j\pi/(m-1))$, $j = 0, 1, \dots, m-1$, with weights $1/(m-1)$, if $1 \le j \le m-2$, and $1/\{2(m-1)\}$, if $j = 0$ or $m-1$. Now, let $\alpha_1 > 0$ and $\alpha_2 > 0$. Then, ξ is supported by the maximum points of $4\alpha_1 T_{m-1}^2(x) + \alpha_2(T_m(x) - \beta T_{m-2}(x))^2$, $x \in [-1, 1]$, with weights given implicitly by (28). Calculations, too long to be reported here, lead to the following. If $\alpha_1 < \alpha_2$, then the support points of ξ are the $m + 1$ zeros of

$$(1 - x^2)(U_{m-1}(x) + \beta U_{m-3}(x)),$$

where U_r, $r = 0, 1, \dots,$ denotes the rth Tchebycheff polynomial of the second kind, and $U_{-1} = 0$. The weights are

$$\xi(x_j) = (1 - \beta^2)/\left\{(m-1)(1-\beta^2) + (1+\beta)^2 - 4\beta T_{m-1}^2(x_j)\right\},$$

for $j = 1, \dots, m-1$, where $x_1, \dots, x_{m-1}$ denote the support points in the interior of $[-1, 1]$, and

$$\xi(1) = \xi(-1) = \tfrac{1}{2}(1 - \beta^2)/\left\{(m-1)(1-\beta^2) - (1-\beta)^2\right\}.$$

Let us denote this design by ξ_β [where $\beta = \alpha_1/\alpha_2 \in (0, 1)$]. If $\alpha_1 \ge \alpha_2$, then one obtains $\xi = \xi_1$. From Theorems 2 and 2(a) we conclude, that the family of designs ξ_β, $0 \le \beta \le 1$, constitutes the class of all $\mathscr{Q}$-invariant, admissible designs for (a_{m-1}, a_m). Note that ξ_0 and ξ_1 are admissible for (a_{m-1}, a_m), since they are the unique designs which are optimal for estimating a_m or a_{m-1}, respectively. Next, we will find ϕ_p-optimal designs for $(a_{m-1}, a_m)'$, $-1 \le p \le \infty$. By Theorems 1 and 1(a), and by the above, a $\mathscr{Q}$-invariant design is ϕ_p-optimal, if and only if $\xi = \xi_\beta$, $0 \le \beta < 1$, $\beta = \alpha_1/\alpha_2$ and $-\mathrm{diag}(\alpha_1, \alpha_2)$ is a subgradient of ϕ_p at $C = \mathrm{diag}(c_1(\beta), c_2(\beta))$, where

$$c_1(\beta) = \int_{[-1,1]} \left\{2^{-(m-2)}T_{m-1}\right\}^2 d\xi_\beta,$$

$$c_2(\beta) = \int_{[-1,1]} \left\{2^{-(m-1)}(T_m - \beta T_{m-2})\right\}^2 d\xi_\beta.$$

Calculations yield

$$c_1(\beta) = 2^{-2m+3}(1+\beta), \qquad c_2(\beta) = 2^{-2m+2}(1-\beta^2).$$

If $p < \infty$, then the gradient of ϕ_p at C is a negative multiple of $\mathrm{diag}(c_1(\beta)^{-p-1}, c_2(\beta)^{-p-1})$, and hence the ϕ_p-optimal design is ξ_β with $\beta = \{c_1(\beta)/c_2(\beta)\}^{-p-1}$, or, equivalently,

$$\left(\frac{1-\beta}{2}\right)^{p+1} - \beta = 0, \qquad 0 \le \beta < 1.$$

There is a unique solution $\beta = \beta_p$, if $p > -1$, and no solution, if $p = -1$ (and hence no ϕ_{-1}-optimal design). Now let $p = \infty$. Since $c_1(\beta) > c_2(\beta)$ for all $\beta \in [0,1)$, the gradient of ϕ_∞ at C exists, and is a negative multiple of diag(0, 1). Hence $\beta_\infty = 0$, and ξ_0 is E-optimal for $(a_{m-1}, a_m)'$.

Two cases of the above were solved in previous work. For $m = 2$, the ϕ_p-optimal designs, $-1 < p \le \infty$, were given by Pukelsheim (1980), Examples 6.2.4 and 6.2.5, which are supported by $-1, 0, 1$ with weights $(1+\beta_p)/4$, $(1-\beta_p)/2$, $(1+\beta_p)/4$. For $p = 0$ and m arbitrary, Studden (1980a) found the D-optimal designs for the s highest coefficients of (26). In case $s = 2$, as considered in our example, we have $\beta_0 = 1/3$, and our D-optimal design $\xi_{1/3}$ coincides with that of Studden (Theorems 4.1 and 4.2 of his paper), which can be seen by using some identities on Tchebycheff polynomials [formulas 22.7.5, 22.5.8 and 22.7.27 in Abramowitz and Stegun (1964)].

REMARK 5. Nonexistence of ϕ_{-1}-optimal designs, as in Examples 1 and 3, is due to our requirement, that under an optimal design ξ_0 the parameters of interest should be estimable [or, equivalently, $J(\xi_0)$ from (2) is nonsingular]. If we drop this requirement, then a ϕ_{-1}-optimal design would be any design ξ_0, which minimizes

$$\phi_{-1}(J(\xi)) = s/\{\mathrm{tr}(J(\xi))\},$$

over the set of all designs ξ (where $s/0$ is defined to be ∞). By (2),

$$\phi_{-1}(J(\xi)) = \max_L \{s/\mathrm{tr}(LM(\xi)L')\},$$

which is a lower semicontinuous function of $M(\xi)$. By compactness of the set of all $M(\xi)$ a ϕ_{-1}-optimal design ξ_0 with possibly singular $J(\xi_0)$ exists.

Acknowledgments. The author wishes to thank the three referees and an Associate Editor for many helpful suggestions and remarks, which led to an improvement of an earlier version of this paper.

REFERENCES

ABRAMOWITZ, M. and STEGUN, I. A. (eds.) (1964). *Handbook of Mathematical Functions. Applied Mathematics Series* **55**. National Bureau of Standards. U.S. Dept. of Commerce, Washington.

ATWOOD, C. L. (1969). Optimal and efficient designs of experiments. *Ann. Math. Statist.* **40** 1570–1602.

GAFFKE, N. (1985a). Directional derivatives of optimality criteria at singular matrices in convex design theory. *Statistics* **16** 373–388.

GAFFKE, N. (1985b). Singular information matrices, directional derivatives, and subgradients in optimal design theory. In *Linear Statistical Inference. Proc. Internat. Conf., Poznan (Poland), 1984* (T. Calinski, and W. Klonecki, eds.) 61–77. Springer, Berlin.

GAFFKE, N. and KRAFFT, O. (1982). Matrix inequalities in the Löwner ordering. In *Modern Applied Mathematics, Optimization and Operations Research* (B. Korte, ed.) 595–622. North-Holland, Amsterdam.

KARLIN, S. and STUDDEN, W. J. (1966). Optimal experimental designs. *Ann. Math. Statist.* **37** 783–815.

KIEFER, J. (1959). Optimum experimental designs. *J. Roy. Statist. Soc. Ser. B* **21** 272–304.

KIEFER, J. (1961). Optimum designs in regression problems. II. *Ann. Math. Statist.* **32** 298–325.

KIEFER, J. (1974). General equivalence theory for optimum designs (approximate theory). *Ann. Statist.* **2** 849–879.

KIEFER, J. (1975). Construction and optimality of generalized Youden designs. In *A Survey of Statistical Design and Linear Models* (J. N. Srivastava, ed.) 333–353. North-Holland, Amsterdam.

KIEFER, J. and WOLFOWITZ, J. (1959). Optimum designs in regression problems. *Ann. Math. Statist.* **30** 271–294.

KRAFFT, O. (1983). A matrix optimization problem. *Linear Algebra Appl.* **51** 137–142.

MÜLLER-FUNK, U., PUKELSHEIM, F. and WITTING, H. (1985). On the duality between locally optimal tests and optimal experimental designs. *Linear Algebra Appl.* **67** 19–34.

NÄTHER, W. and REINSCH, V. (1981). D_s-optimality and Whittle's equivalence theorem. *Math. Operationsforsch. Statist. Ser. Statist.* **12** 307–316.

PARTHASARATHY, T. and RAGHAVAN, T. E. S. (1971). *Some Topics in Two Person Games*. Elsevier, New York.

PUKELSHEIM, F. (1980). On linear regression designs which maximize information. *J. Statist. Plann. Inference* **4** 339–364.

PUKELSHEIM, F. (1987). Information increasing orderings in experimental design theory. *Internat. Statist. Rev.* **55.** To appear.

PUKELSHEIM, F. and TITTERINGTON, D. M. (1983). General differential and Lagrangian theory for optimal experimental design. *Ann. Statist.* **11** 1060–1068.

RAO, C. R. and MITRA, S. K. (1971). *Generalized Inverse of Matrices and Its Applications*. Wiley, New York.

SIBSON, R. (1974). D_A-optimality and duality. In *Progress in Statistics, European Meeting of Statisticians. Colloquia Mathematica Soc. J. Bolyai* **9** (J. Gani et al., eds.) 677–692. North-Holland, Amsterdam.

STEPNIAK, C., WANG, S.-G. and WU, C. F. J. (1984). Comparison of linear experiments with known covariances. *Ann. Statist.* **12** 358–365.

STUDDEN, W. J. (1980a). D_s-optimal designs for polynomial regression using continued fractions. *Ann. Statist.* **8** 1132–1141.

STUDDEN, W. J. (1980b). On a problem of Chebyshev. *J. Approx. Theory* **29** 253–260.

MATHEMATICAL INSTITUTE
UNIVERSITY OF AUGSBURG
MEMMINGER STRASSE 6
D-8900 AUGSBURG
FEDERAL REPUBLIC OF GERMANY