The Annals of Statistics
1995, Vol. 23, No. 1, 30–40

MINIMAX DESIGNS IN LINEAR REGRESSION MODELS

BY H. DETTE,[1] B. HEILIGERS AND W. J. STUDDEN[2]

Universität Göttingen, Universität Augsburg and Purdue University

In the usual linear regression model we investigate the geometric structure of a class of minimax optimality criteria containing Elfving's minimax and Kiefer's ϕ_p-criteria as special cases. It is shown that the optimal designs with respect to these criteria are also optimal for $A'\theta$, where A is any inball vector (in an appropriate norm) of a generalized Elfving set. The results explain the particular role of the A- and E-optimality criterion and are applied for determining the optimal design with respect to Elfving's minimax criterion in polynomial regression up to degree 9.

1. Introduction. For a compact metric space $\mathscr{X}$ which contains at least k different points we consider the usual linear regression model $y = f(x)'\theta$, $x \in \mathscr{X}$. For each $x \in \mathscr{X}$ a random variable $Y(x)$ with mean $f(x)'\theta$ and variance $\sigma^2 > 0$ can be observed, where different observations are assumed to be uncorrelated. The vector of continuous, real-valued and linearly independent regression functions $f(x) = (f_1(x), \dots, f_k(x))'$ is known, while $\theta \in \mathbb{R}^k$ is an unknown parameter vector. A design ξ is a probability measure on a sigma field on $\mathscr{X}$ which contains all one-point sets. The performance of a given design is evaluated by its information matrix

$$M(\xi) = \int_{\mathscr{X}} f(x)f(x)'\, d\xi(x) \in \mathbb{R}^{k\times k}.$$

If ξ is an exact design concentrating masses n_i/n at the points x_i, $i = 1, \dots, s$, the information matrix $M(\xi)$ is proportional to the inverse of the covariance matrix of the least squares estimator calculated from n observations, n_i at x_i, $i = 1, \dots, s$.

Almost all optimality criteria which can be used to discriminate between competing designs depend on the information matrix $M(\xi)$ or its inverse [see, e.g., Silvey (1980) or Pukelsheim (1993)]. In this paper we will consider the geometric structure of two generalizations of the E-optimality criterion which minimizes the maximum eigenvalue of the inverse of the information matrix. The first extension of this criterion is due to Kiefer [(1974), equation (4.18); see also Kiefer (1975), page 337], who defines a design ξ_p to be ϕ_p-optimal if ξ_p minimizes

$$\phi_p(M(\xi)) = \begin{cases} (\operatorname{tr}(M(\xi)^{-p}))^{1/p}, & \text{if } M(\xi) \text{ is positive definite,} \\ \infty, & \text{otherwise.} \end{cases} \tag{1.1}$$

Received September 1992; revised February 1994.
[1]Research supported in part by the Deutsche Forschungsgemeinschaft.
[2]Research supported in part by NSF Grant DMS-91-01730.
AMS 1991 *subject classification*. Primary 62K05.
Key words and phrases. Approximate design theory, minimax criterion, Elfving set, polynomial regression.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Statistics.* www.jstor.org

Here, $1 \leq p \leq \infty$ and the case $p = \infty$ gives the E-optimality criterion. Note that we have omitted the factor $1/k$ in our definition and that Kiefer's ϕ_p-criteria can be also considered for the case $-1 \leq p < 1$ [see Pukelsheim (1980)], but throughout this paper we will assume that $p \geq 1$. A generalization of the E-optimality criterion in a different direction results from the Courant Fischer characterization of the maximum eigenvalue of $M^{-1}(\xi)$,

$$\lambda_{\max}(M^{-1}(\xi)) = \max\{c'M^{-1}(\xi)c \mid c \in \mathbb{R}^k,\ |c|_2 = 1\}$$

(here $|\cdot|_2$ denotes the Euclidean norm on $\mathbb{R}^k$). Replacing the Euclidean norm $|\cdot|_2$ by an arbitrary norm $|\cdot|$ on $\mathbb{R}^k$, we will call a design ξ minimax optimal with respect to the $|\cdot|$-norm if ξ minimizes

$$\phi_{|\cdot|}(M(\xi)) = \max\{c'M^{-1}(\xi)c \mid c \in \mathbb{R}^k,\ |c| = 1\}. \tag{1.2}$$

We will omit the dependency on the norm in this definition whenever it is clear from the context which norm is used in the minimax optimality criterion (1.2).

In Section 2 we introduce a general minimax criterion which contains (1.1) and (1.2) as special cases. It is shown that the minimax optimal design with respect to this criterion is also optimal for $A'\theta$, where $A \in \mathbb{R}^{k\times k^2}$ is an inball vector of a k^3-dimensional Elfving set (in an appropriate norm). The criteria (1.1) and (1.2) are discussed as special cases in Section 3. Finally, the results are applied in Section 4 for the determination of the optimal design with respect to Elfving's minimax criterion [Elfving (1959)] in polynomial regression models up to degree 9.

2. Optimal minimax designs. Let $l \in \mathbb{N}$ and let $|\cdot|$ denote an arbitrary matrix norm on $\mathbb{R}^{k\times l}$ with dual or conjugate norm $|\cdot|_*$, that is,

$$|D|_* := \max\{\operatorname{tr}(D'C) \mid C \in \mathbb{R}^{k\times l},\ |C| = 1\} \tag{2.1}$$

[see, e.g., von Neumann (1937), Rockafellar (1970) or Zietak (1988)]. The unit spheres of $|\cdot|$ and $|\cdot|_*$ are denoted by $\mathscr{C}$ and $\mathscr{D}_*$, respectively, and we define a minimax criterion $\phi_{\mathscr{C}}$ and an information function $j_{\mathscr{D}_*}$ by

$$\begin{aligned} \phi_{\mathscr{C}}(M(\xi)) &= \max\{\operatorname{tr}(C'M^{-1}(\xi)C) \mid C \in \mathscr{C}\}, \qquad M(\xi) > 0,\\ j_{\mathscr{D}_*}(M(\xi)) &= \min\{\operatorname{tr}(D'M(\xi)D) \mid D \in \mathscr{D}_*\}, \qquad M(\xi) \geq 0. \end{aligned} \tag{2.2}$$

A design is called minimax optimal (with respect to the norm $|\cdot|$) if it minimizes $\phi_{\mathscr{C}}(M(\xi))$. In the following we will need an equivalence theorem for minimax optimal designs which can easily be obtained from general equivalence theorems for optimal designs [see, e.g., Gaffke (1985, 1987), Pukelsheim (1993), or Hoang and Seeger (1991)].

PROPOSITION 2.1. *A design ξ_M is minimax optimal with respect to the $|\cdot|$-norm if and only if there exist an integer $1 \leq k_0 \leq k$, matrices $D_1, \ldots, D_{k_0} \in \mathscr{D}_*$*

and positive numbers $\alpha_1, \dots, \alpha_{k_0}$ *with* $\sum_{i=1}^{k_0} \alpha_i = 1$ *such that* $\operatorname{tr}(D_i' M(\xi_M) D_i) = j_{\mathscr{D}_*}(M(\xi_M))$, $i = 1, \dots, k_0$, *and*

$$\sum_{i=1}^{k_0} \alpha_i \operatorname{tr}(D_i' f(x) f(x)' D_i) \le j_{\mathscr{D}_*}(M(\xi_M)) \quad \textit{for all } x \in \mathscr{X}. \tag{2.3}$$

Moreover, in that case the quantities $\operatorname{tr}(D_i' M(\xi_M) D_i)$ $(i = 1, \dots, k_0)$ *and* $j_{\mathscr{D}_*}(M(\xi_M))$ *do not depend on the choice of the minimax design* ξ_M.

LEMMA 2.2. *Let* $M > 0$. *Then* $\phi_{\mathscr{C}}(M) = [j_{\mathscr{D}_*}(M)]^{-1}$. *Moreover* $C_0 \in \mathscr{C}$ *maximizes* $\operatorname{tr}(C' M^{-1} C)$ *over* $\mathscr{C}$ *if and only if* $D_0 = j_{\mathscr{D}_*}(M) M^{-1} C_0$ *is an element of* $\mathscr{D}_*$ *and minimizes* $\operatorname{tr}(D' M D)$.

PROOF. The relation $\phi_{\mathscr{C}}(M) = [j_{\mathscr{D}_*}(M)]^{-1}$ follows from Cauchy's inequality. If $C_0 \in \mathscr{C}$ maximizes $\operatorname{tr}(C' M^{-1} C)$, we have, for all $C \in \mathscr{C}$,

$$\operatorname{tr}^2(C' D_0) \le (j_{\mathscr{D}_*}(M))^2 \operatorname{tr}(C' M^{-1} C) \operatorname{tr}(C_0' M^{-1} C_0) \le 1$$

(with equality for $C = C_0$), which shows that $D_0 \in \mathscr{D}_*$. Conversely, if $D_0 \in \mathscr{D}_*$ minimizes $\operatorname{tr}(D' M D)$ and $C_0 \in \mathscr{C}$ satisfies $\operatorname{tr}(C_0' D_0) = |D_0|_* = 1$, then

$$1 = \operatorname{tr}^2(C_0' D_0) \le \operatorname{tr}(C_0' M^{-1} C_0) \operatorname{tr}(D_0' M D_0) \le j_{\mathscr{D}_*}(M) \phi_{\mathscr{C}}(M) = 1,$$

which shows that C_0 maximizes $\operatorname{tr}(C' M^{-1} C)$ and $C_0 = [j_{\mathscr{D}_*}(M)]^{-1} M D_0$. □

REMARK 2.3. By Proposition 2.1 the minimax optimal design problem is related to the nonlinear approximation problem of Chebyshev-type

$$\text{minimize} \max_{x \in \mathscr{X}} f'(x) E f(x) \quad \text{over } E \in \operatorname{co}(\{DD' \mid |D \in \mathscr{D}_*\}) \ (= \mathscr{E}, \text{ say})$$

($\operatorname{co}(\mathscr{A})$ denotes the convex hull of a set $\mathscr{A}$), which is similar to that considered in Heiligers (1994) (for weighted polynomial regression and ϕ given by (1.2)). Due to the complicated structure of the feasibility set $\mathscr{E}$, in general there is only small hope to find explicit solutions. The problem, however, substantially simplifies if $l = 1$ and if the number k_0 from Proposition 2.1 is known to be 1, since then it is equivalent to the linear approximation problem of Chebyshev-type

$$\text{minimize} \max_{x \in \mathscr{X}} |D' f(x)| \quad \text{over } D \in \mathscr{D}_*,$$

(see also Section 4).

We remark that for all $E = \sum_i \alpha_i D_i D_i' \in \mathscr{E}$ with $\alpha_i > 0$ for all i, and all positive definite matrices M such that $tr(EM) = j_{\mathscr{D}_*}(M)$, the matrix $-\phi_{\mathscr{C}}(M) E$ is a subgradient of $\log(\phi_{\mathscr{C}})$ at M [see Gaffke (1985), Lemma 3].

Throughout this paper we will use the following matrix norm on $\mathbb{R}^{k \times lm}$, $m \in \mathbb{N}$, induced by a given vector norm $|\cdot|$ on $\mathbb{R}^{k \times l}$. For a given matrix $\tilde{A} = (A_1, \dots, A_m) \in \mathbb{R}^{k \times lm}$, $A_i \in \mathbb{R}^{k \times l}$, define

$$||A|| = \left(\sum_{i=1}^{m} |A_i|^2 \right)^{1/2}.$$

It is easy to see that the dual norm of $||\cdot||$ is given by

$$(2.4)\qquad ||\tilde{D}||_* = \left(\sum_{i=1}^{m} |D_i|_*^2\right)^{1/2},$$

where $|\cdot|_*$ is the dual of the given matrix norm $|\cdot|$ on $\mathbb{R}^{k\times l}$, $\tilde{D} = (D_1, \dots, D_m)$. We consider a generalized Elfving set

$$(2.5)\qquad \begin{aligned} \mathscr{R}_m^{(l)} &= \mathrm{co}\left(\left\{(f(x)\varepsilon_1', \dots, f(x)\varepsilon_m') \,\middle|\, x \in \mathscr{X},\ \varepsilon_j \in \mathbb{R}^l,\ \sum_{j=1}^{m} |\varepsilon_j|_2^2 = 1\right\}\right) \\ &\subseteq \mathbb{R}^{k\times lm}. \end{aligned}$$

Note that $\mathscr{R}_m^{(l)}$ is convex, compact, symmetric with respect to the origin and that for $l = m = 1$ this definition gives the set introduced by Elfving (1952), while for $l = 1$ or $m = 1$ the definition (2.5) yields the generalized Elfving set considered in Studden (1971). A more general version of this set and some examples illustrating its geometric structure are discussed in the context of model robust designs by Dette (1993). The minimum distance of all boundary points of $\mathscr{R}_m^{(l)}$ to the origin

$$r_m^{(l)} = \min\{||\tilde{A}|| \mid \tilde{A} \in \partial\mathscr{R}_m^{(l)}\}$$

is called the inball radius of $\mathscr{R}_m^{(l)}$, and every matrix $\tilde{A}$ with $||\tilde{A}|| = r_m^{(l)}$ is called an inball vector of $\mathscr{R}_m^{(l)}$. The following theorem shows that inball radii and vectors of the Elfving set in (2.5) are intimately related to the minimax optimal design problem.

THEOREM 2.4. *Let $m \geq k_0$ and let $\alpha_1, \dots, \alpha_{k_0}$ and $D_1, \dots, D_{k_0} \in \mathscr{D}_*$ denote the quantities from Proposition* 2.1.

(a) *Let $\tilde{D} = (j_{\mathscr{D}_*}(M(\xi_M)))^{-1/2}$, $(\sqrt{\alpha_1}D_1, \dots, \sqrt{\alpha_{k_0}}D_{k_0}, 0, \dots, 0) \in \mathbb{R}^{k\times lm}$, and define $\tilde{A} = M(\xi_M)\tilde{D}$. Then $\tilde{A}$ is a $||\cdot||$-inball vector of $\mathscr{R}_m^{(l)}$ with supporting hyperplane $\tilde{D}$. The $||\cdot||$-inball radius is given by $r_m^{(l)} = (\phi_{\mathscr{E}}(M(\xi_M)))^{-1/2}$.*

(b) *The minimax optimal design ξ_M (with respect to the $|\cdot|$-norm) is optimal for $\tilde{A}'\theta$, where $\tilde{A} \in \mathbb{R}^{k\times lm}$ is any $||\cdot||$-inball vector of $\mathscr{R}_m^{(l)}$. If $\tilde{D} \in \mathbb{R}^{k\times lm}$ is a supporting hyperplane to $\mathscr{R}_m^{(l)}$ at the $||\cdot||$-inball vector $\tilde{A}$, we have $|\tilde{D}'f(x_i)|_2 = 1$ for all support points x_i of ξ_M.*

PROOF. Let $\tilde{N} = (N_1, \dots, N_m) \in \mathbb{R}^{k\times lm}$, $N_i \neq 0$. Then we have, for all $k \times k$ matrices $B \geq 0$,

$$j_{\mathscr{D}_*}(B) = \min\left\{\frac{\mathrm{tr}(N'BN)}{|N|_*^2} \,\middle|\, N \in \mathbb{R}^{k\times l}\setminus\{0\}\right\} \leq \frac{\mathrm{tr}(N_i'BN_i)}{|N_i|_*^2}, \qquad i = 1, \dots, m,$$

which implies [using (2.4)]

$$||\tilde{N}||_*^2 \leq \frac{\sum_{i=1}^{m}\mathrm{tr}(N_iN_i'B)}{j_{\mathscr{D}_*}(B)} = \frac{\mathrm{tr}(\tilde{N}\tilde{N}'B)}{j_{\mathscr{D}_*}(B)}.$$

Because $j_{\mathscr{D}_*}$ is an information function [see Pukelsheim (1980)] it thus follows, for the polar function of $j_{\mathscr{D}_*}$,

$$j^o_{\mathscr{D}_*}(\tilde{N}\tilde{N}') = \inf\left\{ \frac{\operatorname{tr}(\tilde{N}\tilde{N}'B)}{j_{\mathscr{D}_*}(B)} \,\Big|\, B \neq 0 \right\} \geq ||\tilde{N}||_*^2. \tag{2.6}$$

From the definition of $\tilde{D}$ and $\tilde{A}$ we have $\operatorname{tr}(\tilde{D}'\tilde{A}) = 1$, and Proposition 2.1 implies that $\tilde{A} \in \partial\mathscr{R}_m^{(l)}$ with supporting hyperplane $\tilde{D}$. Moreover, Lemma 2.2, (2.6) and Pukelsheim's "mutual boundedness" Theorem 3 [see Pukelsheim (1980)] yield that

$$\begin{aligned} \left[r_m^{(l)}\right]^2 \leq ||\tilde{A}||^2 = [\phi_{\mathscr{C}}(M(\xi_M))]^{-1} = j_{\mathscr{D}_*}(M(\xi_M)) \\ \leq \frac{1}{j^o_{\mathscr{D}_*}(\tilde{N}\tilde{N}')} \leq \frac{1}{||\tilde{N}||_*^2}, \end{aligned} \tag{2.7}$$

for all covering half-spaces $\tilde{N}$ of $\mathscr{R}_m^{(l)}$ [i.e., $|\tilde{N}f(x)|_2^2 = \sum_{i=1}^m f(x)'N_iN_i'f(x) \leq 1 \ \forall\, x \in \mathscr{X}$]. Thus, using the representation

$$r_m^{(l)} = \min\left\{ \frac{1}{||\tilde{N}||_*} \,\Big|\, \tilde{N} \in \mathbb{R}^{k\times lm},\ |\tilde{N}'f(x)|_2 \leq 1,\ \forall\, x \in \mathscr{X} \right\}, \tag{2.8}$$

the assertion (a) follows. Part (b) is proved by exactly the same arguments as in Dette and Studden (1993), and the proof is therefore omitted. □

REMARK 2.5. If $\tilde{D} = (D_1, \ldots, D_m) \in \mathbb{R}^{k\times lm}$ is a covering half-space to $\mathscr{R}_m^{(l)}$ achieving the minimum in (2.8), then the matrix $\tilde{A} = (|D_1|_*A_1, \ldots, |D_m|_*A_m)/||D||_*^2$ defines a $||\cdot||$-inball vector of the Elfving set $\mathscr{R}_m^{(l)}$, where $A_j \in \mathbb{R}^{k\times l}$ is any matrix satisfying

$$|A_j| = 1, \qquad \operatorname{tr}(D_j'A_j) = |D_j|_*, \qquad j = 1, \ldots, m.$$

[The matrix A_j is called dual of D_j with respect to the $|\cdot|$-norm; see Zietak (1988).] Even if the optimal covering half-space cannot be determined, the covering half-spaces of $\mathscr{R}_m^{(l)}$ provide lower bounds for the minimax efficiency

$$\operatorname{Eff}_{\mathscr{C}}(\xi) := \frac{\phi_{\mathscr{C}}(M(\xi_M))}{\phi_{\mathscr{C}}(M(\xi))}$$

of a given design ξ when the optimal minimax design ξ_M with respect to the $|\cdot|$-norm is unknown.

COROLLARY 2.6. *Let $m \geq 1$ and let $\tilde{D}$ denote a supporting hyperplane to $\mathscr{R}_m^{(l)}$. Then the minimax efficiency (with respect to the $|\cdot|$-norm) of a given design ξ is bounded by*

$$\operatorname{Eff}_{\mathscr{C}}(\xi) \geq \frac{j^o_{\mathscr{D}_*}(\tilde{D}\tilde{D}')}{\phi_{\mathscr{C}}(M(\xi))} \geq \frac{||\tilde{D}||_*^2}{\phi_{\mathscr{C}}(M(\xi))}.$$

If $\tilde{D}$ is an optimal supporting hyperplane [i.e., $\tilde{D}$ minimizes (2.8)], then the equality $j^o_{\mathscr{D}_*}(\tilde{D}\tilde{D}') = ||\tilde{D}||_*^2$ holds true.

PROOF. This is an immediate consequence of (2.7) and (2.8). □

REMARK 2.7. The results of Theorem 2.4 can easily be generalized to minimax optimal design problems for parameter subsystems. For a given $k \times s$ matrix K of rank s a minimax optimal design for $K'\theta$ allows the estimability of $K'\theta$ [i.e., range$(K) \subseteq$ range$(M(\xi))$] and minimizes $\phi_{\mathscr{C}}((K'M(\xi)^{-}K)^{-1})$. According to Gaffke [(1987), Theorem 1] there exists a left inverse $L_0' \in \mathbb{R}^{s\times k}$ of K such that the minimax optimal design for $K'\theta$ is minimax optimal for the full parameter vector in the "new" regression setup $y = \tilde{\theta}'\tilde{f}(x)$, where $\tilde{f}(x) = L_0'f(x)$. Thus we obtain from Theorem 2.4 that the minimax optimal design for $K'\theta$ is optimal for $A'\tilde{\theta}$ for any $\|\cdot\|$-inball vector $A \in \mathbb{R}^{s\times sl}$ of the Elfving set $\tilde{\mathscr{R}}_s^{(l)}$, where $\mathscr{R}_m^{(l)}$ is defined as

$$\tilde{\mathscr{R}}_m^{(l)} = \mathrm{co}\left(\left\{L_0'f(x)(\varepsilon_1',\ldots,\varepsilon_m')\,\middle|\, x \in \mathscr{X},\ \varepsilon_j \in \mathbb{R}^l,\ \sum_{j=1}^m |\varepsilon_j|_2^2 = 1\right\}\right) \subseteq \mathbb{R}^{s\times ml},$$

$m = 1,\ldots,s$. The applicability of this result is limited (except in the case $s = k$, where $L_0' = K^{-1}$) because in general L_0 is unknown and a $\|\cdot\|$-inball vector of $\tilde{\mathscr{R}}_s^{(l)}$ cannot be found.

3. Elfving's minimax and Kiefer's ϕ_p-criterion. In this section we will return to the criteria defined in (1.1) and (1.2), which now emerge as special cases from the general theory of Section 2, with the same Elfving set (2.5) for both criteria.

First, let $l = 1$. Then criterion (2.2) reduces to the minimax criterion (1.2). The geometric structure of the minimax problem is described in Theorem 2.4 ($l = 1$), where the generalized Elfving set in (2.5) reduces to the set

$$\mathscr{R}_m = \mathrm{co}(\{f(x)\varepsilon' \mid x \in \mathscr{X},\ \varepsilon \in \mathbb{R}^m, |\varepsilon|_2 = 1\}), \tag{3.1}$$

which was first introduced by Studden (1971) when characterizing optimal designs for $A'\theta$ (here $A \in \mathbb{R}^{k\times m}$ is a given matrix). Theorem 2.4 now generalizes the results of Dette and Studden (1993) ($|\cdot| = |\cdot|_2$) to arbitrary criteria of the form (1.2). The following important examples are mentioned as special cases.

1. Considering the l_2-norm, we obtain the E-optimality criterion, while the l_1-norm yields to Elfving's minimax criterion [Elfving (1959)], that is,

$$\phi_{|\cdot|_1}(M(\xi)) = \max\{c'M^{-1}(\xi)c \mid |c|_1 = 1\} = \max_{i=1}^{k}\{M^{-1}(\xi)\}_{ii}. \tag{3.2}$$

2. If the regression norm [see Pukelsheim (1981)]

$$|c|^R = \inf\{\alpha \geq 0 \mid c \in \alpha\mathscr{R}_1\}$$

on $\mathbb{R}^k$ is used in definition (1.2), then it is straightforward to see that the optimality criterion (1.2) gives the well known G-optimality criterion, that is,

$$\phi_{|\cdot|^R}(M(\xi)) = \max\{c'M^{-1}(\xi)c \mid c \in \partial\mathscr{R}_1\} = \max_{x\in\mathscr{X}} f(x)'M^{-1}(\xi)f(x)$$

(note that $|\cdot|^R$ characterizes the Elfving set $\mathscr{R}_1$ as the unit ball). The dual norm of $|\cdot|^R$ is given by $|d|_*^R = \max_{x\in\mathscr{X}} |d'f(x)|$ [see, e.g., Householder (1965)].

Second, let $l = k$ and define a norm on $\mathbb{R}^{k\times k}$ by

$$||A||_{p'} = |\sigma(A)|_{p'} = (\mathrm{tr}(AA')^{p'/2})^{1/p'}, \qquad 1 \le p' \le \infty,$$

where $\sigma_1(A) \le \cdots \le \sigma_k(A)$ denote the singular values of a given matrix $A \in \mathbb{R}^{k\times k}$, $\sigma(A) = (\sigma_1(A), \dots, \sigma_k(A))'$ and $|\cdot|_{p'}$ is the $l_{p'}$-norm on $\mathbb{R}^k$. Setting $p' = 2p/(p-1)$ we obtain, for the optimality criterion (2.2),

$$\begin{aligned}\phi_{\mathscr{C}}(M(\xi)) &= \max\{\mathrm{tr}(C'M^{-1}(\xi)C) \mid C \in \mathbb{R}^{k\times k},\ ||C||_{p'} = 1\}\\ &= \max\{\mathrm{tr}(B'M^{-1}(\xi)) \mid B \ge 0,\ ||B||_{p/(p-1)} = 1\}\\ &= ||M^{-1}(\xi)||_p = \phi_p(M(\xi)),\end{aligned}$$

where the last line follows from Gaffke and Krafft [(1982), Theorem 5.10]. Using Pukelsheim [(1980), Lemma 3] we obtain that for $1 \le p < \infty$ the quantities in Proposition 2.1 are given by

$$k_0 = 1 \quad \text{and} \quad D_1 = (\mathrm{tr}(M(\xi_p)^{-p}))^{-(p+1)/2p} M(\xi_p)^{-(p+1)/2} Q, \tag{3.3}$$

where Q denotes an arbitrary orthogonal $k \times k$ matrix and ξ_p the ϕ_p-optimal design. If $p = \infty$, a possible choice for D_1 is the matrix

$$D_1 = \big(\sqrt{\beta_1}z_1, \dots, \sqrt{\beta_{k_1}}z_{k_1}, 0, \dots, 0\big)Q \in \mathbb{R}^{k\times k}, \tag{3.4}$$

where $k_1 \le k$, $\beta_j > 0$, $\sum_{j=1}^{k_1} \beta_j = 1$ and $z_1, \dots, z_{k_1}$ are normalized eigenvectors of $M(\xi_\infty)$ corresponding to its minimum eigenvalue which satisfy inequality (2.3) in Proposition 2.1 for the E-optimality criterion [see also Pukelsheim (1980), Corollary 8.1]. By an application of Theorem 2.4 we thus obtain the following result.

COROLLARY 3.1. *For $1 \le p \le \infty$ let ξ_p denote the ϕ_p-optimal design and let D_1 be defined by* (3.3) *if $1 \le p < \infty$ and by* (3.4) *if $p = \infty$.*

(a) *The matrix $A = \phi_p(M(\xi_p))^{1/2} M(\xi_p) D_1$ defines a $||\cdot||_{2q}$-inball vector of the Elfving set $\mathscr{R}_k$ with supporting hyperplane $\phi_p(M(\xi_p))^{1/2} D_1$, $1/p + 1/q = 1$. The $||\cdot||_{2q}$-inball radius of $\mathscr{R}_k$ is given by $(\phi_p(M(\xi_p)))^{-1/2}$.*

(b) *If $1 \le p \le \infty$ and A is any $||\cdot||_{2q}$-inball vector of the Elfving set $\mathscr{R}_k$, then the ϕ_p-optimal design ξ_p is also optimal for $A'\theta$.*

REMARK 3.2. Let $\tilde{p} = 2p/(p+1)$, and let $D \in \mathbb{R}^{k\times k}$ denote an "optimal" covering half-space, that is, $||D||_{\tilde{p}} = 1/r_k$ with singular value decomposition $D = U\,\mathrm{diag}(\sigma(D))V'$ [where $\mathrm{diag}(x_1, \dots, x_k)$ means a diagonal matrix with diagonal elements $x_1, \dots, x_k$]. Then a $||\cdot||_{2q}$-inball vector can be obtained as follows. Consider a dual vector $\sigma^*(D)$ of $\sigma(D) \in \mathbb{R}^k$ with respect to the ℓ_{2q}-norm [that is, $\sigma^*(D)'\sigma(D) = |\sigma(D)|_{\tilde{p}}$, $|\sigma^*(D)|_{2q} = 1$] and define $A = U\,\mathrm{diag}(\sigma^*(D))V'/||D||_{\tilde{p}}$. Thus we obtain $\mathrm{tr}(D'A) = |\sigma(D)|_{\tilde{p}}/||D||_{\tilde{p}} = 1$ and $||A||_{2q} = 1/||D||_{\tilde{p}}$, which shows that A defines an inball vector of $\mathscr{R}_k$.

For $1 < p < \infty$ the strict convexity of the ℓ_{2q}-norm implies that A is the unique $||\cdot||_{2q}$-inball vector corresponding to D [Zietak (1988), Theorem 3.1, Corollary 4.2].

REMARK 3.3. Recalling the discussion in Remark 2.7, we see that Corollary 3.1 gives new insight into the particular role of the ϕ_1-optimality criterion. Here ($q = \infty$) any $s \times s$ orthogonal matrix Q (appropriately scaled) defines a $||\cdot||_\infty$-vector of $\tilde{\mathscr{R}}_1^{(s)}$ [this follows from Corollary 3.1(a)].

It should also be mentioned that the results of this section can easily be generalized for unitarily invariant norms on $\mathbb{R}^{k\times k}$. These norms are obtained by replacing the ℓ_p-norm in (3.1) by a so-called symmetric gauge function $\psi(\cdot)$ on $\mathbb{R}^k$ which satisfies, in addition to the norm properties, the symmetry assumption

$$\psi((\varepsilon_1 a_{i_1}, \dots, \varepsilon_k a_{i_k})') = \psi((a_1, \dots, a_k)'),$$

for all permutations $a_{i_1}, \dots, a_{i_k}$ of $a_1, \dots, a_k$ and for all $\varepsilon_j = \mp 1$ [see von Neumann (1937), Mudholkar (1966), or Zietak (1988) for more details].

4. Elfving's minimax criterion for polynomial regression. Let $l = 1$, $\mathscr{X} = [-1, 1]$, $f(x) = (1, x, \dots, x^d)'$ and $1 \le p \le \infty$. Thus we are faced with the minimax criterion (1.2) with respect to the ℓ_p-norm defined in (3.2). Contrary to an example for spring balance weighing designs ($p = 2$) discussed in Dette and Studden (1993), the situation here is more complicated because we are not able to find the $||\cdot||_p$-inball radius of the Elfving set $\mathscr{R}_{d+1}$ defined in (3.1). However, if the (unknown) number k_0 in Proposition 2.1 is 1, Theorem 2.4(b) shows that the minimax design $\xi_{|\cdot|}$ is already optimal for any $||\cdot||_p$-inball vector c of the first Elfving set $\mathscr{R}_1$. This fact was used by Pukelsheim and Studden (1993) to show that the E-optimal design (minimax with respect to the ℓ_2-norm) is supported at the Chebyshev points $s_j = \cos([(d-j)/d]\pi)$, $j = 0, \dots, d$. Observing these results and Corollary 2.6 it will therefore be useful to find (at least) the $||\cdot||_p$-inball vectors of $\mathscr{R}_1$ and the corresponding optimal designs. The optimal designs for these inball vectors seem to be good candidates for minimax optimality. Throughout this example let ξ_k denote the optimal design minimizing the variance of the least squares estimator for the individual coefficient θ_k in the polynomial regression $y = \theta_0 + \theta_1 x + \dots + \theta_d x^d$ [see Studden (1968)] and define $t = (t_0, \dots, t_d)')$ as the vector of the coefficients of the Chebyshev polynomial of the first kind, that is, $t'f(x) = T_d(x) = \cos(d \arccos x)$.

THEOREM 4.1.

(a) *If* $1 < p < \infty$, *the* $||\cdot||_p$*-inball vector* $c = (c_0, \dots, c_d)'$ *of* $\mathscr{R}_1$ *has coordinates*

$$c_i = \frac{\operatorname{sign}(t_i)|t_i|^{q-1}}{|t|_q^q}, \qquad i = 0, \dots, d.$$

The c-optimal design for this inball vector is given by $\xi_c = \sum_{j=0}^{d} |t_j|^q/|t|_q^q \cdot \xi_j$ *and the* $\|\cdot\|_p$*-inball radius is* $1/|t|_q$.

(b) *If* $p = 1$, *the* $\|\cdot\|_1$*-inball vector has coordinates*

$$c_i = \begin{cases} \operatorname{sign}(t_i)\dfrac{g_i}{|t|_\infty}, & \text{if } |t_i| = |t|_\infty, \\ 0, & \text{if } |t_i| < |t|_\infty, \end{cases}$$

where $g_i \geq 0$ *and* $\Sigma g_i = 1$. *The c-optimal design is given by* $\xi_c = \sum_j g_j \xi_j$ *and the* $\|\cdot\|_1$*-inball radius is* $1/|t|_\infty$.

PROOF. Using (2.8) (for $m = l = 1$), we have to maximize $|a|_q$ subject to the restriction $|a' f(x)| \leq 1$ for all $x \in [-1, 1]$, $a = (a_0, \dots, a_d)' \in \mathbb{R}^{d+1}$. Using a result of Cantor (1977), we obtain, for the coefficients of the vector a,

$$|a_{d-2m}| + |a_{d-2m-1}| \leq |t_{d-2m}|, \qquad m = 0, \dots, \left\lfloor \frac{d-1}{2} \right\rfloor,$$

with equality if and only if $a = \mp t$. This implies $|a|_q \leq |t|_q$, and (2.8) shows that the $\|\cdot\|_p$-inball radius of $\mathscr{R}_1$ is given by $|t|_q^{-1}$. By the discussion in Remark 2.5 we have to find a dual vector of t (with respect to the ℓ_p-norm) which can easily be obtained considering equality in the Hölder inequality [see, e.g., Zietak (1988), page 60]. Thus the assertion about the inball vectors follows directly from Remark 2.5. Let $L_v(x) = \ell_{vo} + \ell_{v1}x + \cdots + \ell_{vd}x^d$ denote the v-th Lagrange interpolation polynomial at the points $s_0, \dots, s_d$. Then it follows from the results of Studden (1968) that the optimal design ξ_{d-2j} for estimating θ_{d-2j} puts masses $|\ell_{vd-2j}|/|t_{d-2j}|$ at the points s_v, $v = 0, \dots, d$, and Elfving's theorem [Elfving (1952)] yields

$$\frac{1}{|t_{d-2j}|} e_{d-2j} = \sum_{v=0}^{d} (-1)^{d-v+j} \frac{|\ell_{vd-2j}|}{|t_{d-2j}|} f(s_v), \qquad j = 0, \dots, \left\lfloor \frac{d-1}{2} \right\rfloor.$$

Expressing the inball vector c as a linear combination of the unit vectors e_{d-2j}, the assertion now follows directly by a further application of Elfving's theorem. □

To be more explicit, consider the case $d = 2$. Then it is straightforward to show that the ℓ_p-optimal design ξ_c puts masses $2^{q-2}/(1+2^q)$ at the points -1 and 1 and mass $(1+2^{q-1})/(1+2^q)$ at the point 0. Using Lagrangian multipliers and Proposition 2.1 it can be shown by tedious computations that ξ_c is in fact the minimax design with respect to the ℓ_p-norm for all $1 \leq p < \infty$.

Recently Pukelsheim and Studden (1993) showed that ξ_c is E-optimal for all $d \in \mathbb{N}$ (i.e., minimax with respect to the ℓ_2-norm). We will conclude with an example demonstrating that this might not be true for arbitrary $p \geq 1$. To this end consider Elfving's minimax criterion (i.e., $p = 1$, $q = \infty$). Using a table of Chebyshev polynomials of the first kind [see, e.g., Davis (1963), page 369] and Theorem 4.1(b), we see that for $d = 1, 2, 3$ the design $\xi_c = \xi_d$ can be considered as a candidate for minimax optimality. For $d = 5, 6, 7, 8, 9$ we get ξ_{d-2} as a minimax candidate while in the case $d = 4$ [note that $T_4(x) = 8x^4 - 8x^2 + 1$]

every convex combination $\alpha\xi_2+(1-\alpha)\xi_4$, $\alpha \in [0,1]$, seems to be a good choice. Tedious algebra and Proposition 2.1 show that for $d = 1,2,3$ the design ξ_d is in fact minimax optimal with respect to Elfving's criterion. In the case $d = 5,6,7,8,9$ the design ξ_{d-2} can be shown to be minimax, while for $d = 4$ every convex combination of ξ_2 and ξ_4 fails to be minimax (with respect to Elfving's criterion). In this case the number k_0 in Proposition 2.1 is 2; the $\|\cdot\|_1$-inball radius of $\mathscr{R}_2$ can only be determined numerically and is smaller than $1/|t|_\infty = \frac{1}{8}$. However, we can use Corollary 2.6 to obtain a lower bound for the minimax efficiency, that is,

$$\mathrm{Eff}_{|\cdot|_1}(\xi) \geq |t|_\infty^2 \cdot \left[\max_{i=0}^{d}(e_i' M^{-1}(\xi)e_i)\right]^{-1}.$$

The average of the optimal designs for the coefficients θ_2 and θ_4, $\xi^* = \frac{1}{2}(\xi_2+\xi_4)$ puts masses $\frac{3}{32}$, $\frac{1}{4}$, $\frac{5}{16}$, $\frac{1}{4}$ and $\frac{3}{32}$ and the points -1, $-1/\sqrt{2}$, 0, $1/\sqrt{2}$ and 1 and has at least minimax efficiency $\mathrm{Eff}_{|\cdot|_1}(\xi^*) \geq \frac{30}{31} \approx 0.9677$, which shows that ξ^* is a good choice with respect to Elfving's minimax criterion. Numerical calculations yield that for $d = 4$ the minimax design is not supported at the Chebyshev points and puts masses 0.0958, 0.246, 0.3164, 0.246 and 0.0958 at the points -1, -0.7086, 0, 0.7086 and 1. Thus the exact minimax efficiency of the design ξ^* is 0.9997.

The results of the last paragraph suggest that for polynomial regression of degree d on the interval $[-1,1]$ the minimax optimal design with respect to Elfving's minimax criterion is specified by the optimal design for the $|\cdot|_1$-inball vector of the first Elfving set $\mathscr{R}_1$ provided that $\#\{j \mid |t_j| = |t|_\infty\} = 1$. A partial proof of this conjecture and a more complete discussion of the problem including minimax optimal designs for parameter subsystems, different design spaces is given in a recent paper of Dette and Studden (1994).

Acknowledgments. This work was done while H. Dette was visiting the Department of Statistics, Purdue University. The authors would like to thank the Department of Statistics for its hospitality and the Deutsche Forschungsgemeinschaft for making this visit possible by its financial support. We are also indebted to an unknown referee for his constructive criticism on an earlier version of this paper.

REFERENCES

CANTOR, D. G. (1977). Solution to problem 6084 [1976, 292]. *Amer. Math. Monthly* **84** 832–833.

DAVIS, P. J. (1963). *Interpolation and Approximation.* Blaisdell, New York.

DETTE, H. (1993). Elfving's theorem for D-optimality. *Ann. Statist.* **21** 753–766.

DETTE, H. and STUDDEN, W. J. (1993). Geometry of E-optimality. *Ann. Statist.* **21** 416–433.

DETTE, H. and STUDDEN, W. J. (1994). Optimal designs with respect to Elfving's partial minimax criterion in polynomial regression. *Ann. Inst. Statist. Math.* **46** 389–403.

ELFVING, G. (1952). Optimum allocation in linear regression theory. *Ann. Math. Statist.* **23** 255–262.

ELFVING G. (1959). Design of linear experiments. In *Probability and Statistics. The Harald Cramér Volume* (U. Grenander, ed.) 58–74. Almgrist & Wiksell, Stockholm.

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

GAFFKE, N. (1987). Further characterizations of design optimality and admissibility for partial parameter estimation. *Ann. Statist.* **15** 942–957.

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

HEILIGERS, B. (1994). E-optimal designs in weighted polynomial regression. *Ann. Statist.* **22** 917–929.

HOANG, T. and SEEGER, A. (1991). On conjugate functions, subgradients and directional derivatives of a class of optimality criteria in experimental design. *Statistics* **22** 349–368.

HOUSEHOLDER, A. S. (1965). *The Theory of Matrices in Numerical Analysis*. Blaisdell, New York.

KIEFER, J. (1974). General equivalence theory for optimum designs. *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.

MUDHOLKAR, G. S. (1966). On confidence bounds associated with multivariate analysis of variance and nonindependence between two sets of variates. *Ann. Math. Statist.* **37** 1736–1746.

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

PUKELSHEIM, F. (1981). On c-optimal design measures. *Math. Operationsforsch. Statist. Ser. Statist.* **12** 13–20.

PUKELSHEIM, F. (1993). *Optimal Design of Experiments*. Wiley, New York.

PUKELSHEIM, F. and STUDDEN, W. J. (1993). E-optimal designs for polynomial regression. *Ann. Statist.* **21** 402–415.

ROCKAFELLAR, R. T. (1970). *Convex Analysis*. Princeton Univ. Press.

SILVEY, S. D. (1980). *Optimal Design*. Chapman and Hall, London.

STUDDEN, W. J. (1968). Optimal designs on Tchebycheff points. *Ann. Math. Statist.* **39** 1435–1447.

STUDDEN, W. J. (1971). Elfving's theorem for quadratic loss. *Ann. Math. Statist.* **42** 1613–1621.

VON NEUMANN, J. (1937). Some matrix inequalities and metrization of metric spaces. *Tomsk. Univ. Rev.* **1** 286–300. [In *Collected Works*, Pergamon, Oxford (1962) 205–219.]

ZIETAK, K. (1988). On the characterization of the extremal points of the unit sphere of matrices. *Linear Algebra Appl.* **106** 57–75.

H. DETTE
INSTITUT FÜR MATHEMATISCHE STOCHASTIK
TECHNISCHE UNIVERSITÄT DRESDEN
MOMMSENSTRASSE 13
01062 DRESDEN
GERMANY

B. HEILIGERS
INSTITUT FÜR MATHEMATIK
UNIVERSITÄT AUGSBURG
UNIVERSITÄTSSTR. 14
86159 AUGSBURG
GERMANY

W. J. STUDDEN
DEPARTMENT OF STATISTICS
PURDUE UNIVERSITY
MATHEMATICAL SCIENCES BUILDING
WEST LAFAYETTE, INDIANA 47907-1399