Werk

Titel: A New Interpretation of Optimality for E-Optimal Designs in linear Regression Mod...

Autor: Dette, Holger

Jahr: 1993

PURL: https://resolver.sub.uni-goettingen.de/purl?358794056_0040|log6

Kontakt/Contact

Digizeitschriften e.V.
SUB Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen

info@digizeitschriften.de

Metrika (1993) 40:37–50

Metrika

A New Interpretation of Optimality for *E*-Optimal Designs in Linear Regression Models

HOLGER DETTE

Institut für Mathematische Stochastik, Universität Göttingen, Lotzestr. 13, 3400 Göttingen, Germany

Abstract: The optimal design problem for the estimation of several linear combinations $c_l'\theta$ ($l = 1, \ldots, m$) is considered in the usual linear regression model $y = f'(x)\theta$ ($f(x) \in \mathbb{R}^k$, $\theta \in \mathbb{R}^k$). An optimal design minimizes a (weighted) p-norm of the variances of the least squares estimates for the different linear combinations $c_l'\theta$. A generalized Elfving theorem is used to derive the relation of the new optimality criterion to the *E*-optimal design problem. It is shown that the *E*-optimal design for the parameter θ minimizes such a (weighted) p-norm whenever the vector $c = (c_1', \ldots, c_k')'$ is an inball vector of a symmetric convex and compact "Elfving set" in $\mathbb{R}^{k^2}$.

1 Introduction

Consider m linear regression models of the form

$$g_l(x) = f_l'(x)\theta_l \quad l = 1, \ldots, m$$

where $f_l'(x) = (f_{l1}(x), \ldots, f_{lk_l}(x))$ is the vector of regression functions ($l = 1, \ldots, m$). The control variable x takes values in a compact space $\mathcal{X}$ with sigma field $\mathcal{B}$ including all one point sets. $\theta_l = (\theta_{l1}, \ldots, \theta_{lk_l})'$ is the vector of unknown parameters and the functions $f_{l_1}(x), \ldots, f_{lk_l}(x)$ are assumed to be real valued and continuous functions on the design space $\mathcal{X}$. The different models are collected in the set

$$\mathcal{F}_m = \{g_l \mid g_l(x) = f_l'(x)\theta_l,\ l = 1, \ldots, m\}\ . \tag{1.1}$$

For every $x \in \mathcal{X}$ a random variable $Y(x)$ with (unknown) mean $g(x) \in \mathcal{F}_m$ and variance $\sigma^2 > 0$ can be observed where different observations are assumed to be

0026–1335/93/1/37–50 $2.50 © 1993 Physica-Verlag, Heidelberg

independent. A design is a probability measure on $\mathcal{X}$ and the information matrix of ξ in the model g_l is defined by

$$M_l(\xi) = \int_{\mathcal{X}} f_l(x) f_l'(x) d\xi(x) \quad l = 1, \ldots, m \ .$$

Throughout this paper we assume that there exists a finite support design ξ on $\mathcal{X}$ such that all information matrices $M_l(\xi)$ are regular. If ξ is supported at s points $x_1, \ldots, x_s$ with masses $\xi(x_i) = \frac{n_i}{n}$ the experimenter takes n_i *uncorrelated observations* at each point x_i $\left(i = 1, \ldots, s, \ \sum_{i=1}^{s} n_i = n \right)$ and the covariance matrix of the least squares estimator for θ_l (in the model $g_l(x)$) is proportional to the inverse of the information matrix $M_l(\xi)$ (see e.g. Krafft (1978)). This paper deals with the E-optimality and the c_l-optimality criterion for the model g_l. An E-optimal design (for the model g_l) maximizes the minimum eigenvalue $\lambda_{\min}(M_l(\xi))$ of the information matrix $M_l(\xi)$ while a c_l-optimal design minimizes $c_l' M_l^-(\xi) c_l$ for a given vector $c_l \in \mathbb{R}^{k_l}$. Here $M_l^-(\xi)$ denotes any generalized inverse of $M_l(\xi)$ and we assume that the vector c_l is estimable by the design ξ, i.e. $c_l \in range(M_l(\xi))$ (see Pukelsheim (1980)). The c_l-optimality is a special case of the following more general definition of optimality. Consider a number $p \in [-\infty, 1]$, m vectors $c_l \in \mathbb{R}^{k_l}$ and numbers $\beta_l > 0$. The vectors c_l are collected in the "grand" vector $(c_1', \ldots, c_m')' \in \mathbb{R}^k$ $\left(k = \sum_{l=1}^{m} k_l \right)$ and the vector $\beta = (\beta_1, \ldots, \beta_m)$ is called a "prior" for the class $\mathcal{F}_m$ in (1.1) where we assume that the components of β sum to one. For $p \in [-\infty, 1] \setminus \{-\infty, 0\}$ define

$$\Phi_{p,\beta}^{c}(\xi) = \left[\sum_{l=1}^{m} \beta_l (c_l' M_l^-(\xi) c_l)^{-p} \right]^{1/p}$$

and let for $p = 0$ and $p = -\infty$

$$\Phi_{0,\beta}^{c}(\xi) = \prod_{l=1}^{m} (c_l' M_l^-(\xi) c_l)^{-\beta_l} \quad \text{and} \quad \Phi_{-\infty,\beta}^{c}(\xi) = \min_{l=1}^{m} (c_l' M_l^-(\xi) c_l)^{-1}$$

respectively. It is easy to see that $\Phi_{p,\beta}^{c}(\xi) = m^{1/p} \Phi_{p,\beta^*}^{\tilde{c}}(\xi)$ where

$$\beta^* = \left(\frac{1}{m}, \ldots, \frac{1}{m} \right) \quad \text{and} \quad \tilde{c} = (\beta_1^{-1/2p} c_1', \ldots, \beta_m^{-1/2p} c_m')' \quad (p \neq 0, -\infty) \ .$$

Definition 1.1: A design ξ for which all linear combinations $c_l'\theta_l$ are estimable is called $\Phi^c_{p,\beta}$-optimal for the class $\mathcal{F}_m$ with respect to the prior β if ξ maximizes the function $\Phi^c_{p,\beta}(\xi)$. If all models in the set $\mathcal{F}_m$ in (1.1) are identical (i.e. $f_1(x) = f_l(x)$, $l = 1, \ldots, m$) the maximizing design is called $\Phi^c_{p,\beta}$-optimal with respect to the prior β.

Note that if $f_l = f_1$ $(l = 1, \ldots, m)$ and $p = -1$ the $\Phi^c_{p,\beta}$-optimality criterion gives the $A'\theta_1$-optimality (defined as in Karlin and Studden (1966)) when putting $A = A_\beta = (\sqrt{\beta_1}c_1, \ldots, \sqrt{\beta_m}c_m) \in \mathbb{R}^{k_1 \times m}$ (see also Studden (1971) for some properties of optimal designs with respect to this criterion).

The Definition 1.1 is motivated by the following point of critism on the "classical" optimal design theory. This theory is based on the fact that the underlying model (namely g_l) is known by the experimenter which is not very realistic from a practical point of view. In scientific experiments it is often the case that a suitable (linear) model for the description of the dependency of Y from x is not known before experiments have been conducted. However, in some cases the experimenter does not know the "true" model exactly but he knows that the model belongs to the given class $\mathcal{F}_m$ defined by (1.1) (e.g. the experimenter is sure that a linear or quadratic polynomial is appropriate). Assume that in every model $g_l \in \mathcal{F}_m$ the linear combinations $c_l'\theta_l$ $(l = 1, \ldots, m)$ (where $c_1, \ldots, c_m$ are given vectors $c_l \in \mathbb{R}^{k_l}$) or different linear combinations in the same model have to be estimated (in this case all vectors f_l would be the same). Because the variances of the least squares estimates for the linear combinations $c_l'\theta_l$ are proportional to $c_l' M_l^-(\xi) c_l$ $(l = 1, \ldots, m)$ an optimal design with respect to one of the criteria of Definition 1.1 allows "precise" estimates of all linear combinations in the given models of $\mathcal{F}_m$. The numbers β_l are used to reflect the experimenters belief about the adequacy of the models g_l or the importance of the linear combinations $c_l'\theta_l$. In this sense the $\Phi^c_{p,\beta}$-optimality criterion can be viewed as a model robust version of the classical c_l-optimality criterion.

Note that the efficiency of a $\Phi^c_{p,\beta}$-optimal design for estimating one linear combination $c_{l_0}'\theta_{l_0}$ in the model g_{l_0} will depend sensitively on the choice of the prior β_{l_0}. Thus a "bad" design for estimating $c_{l_0}'\theta_{l_0}$ in the model g_{l_0} may be judged as a good solution of the $\Phi^c_{p,\beta}$-optimal design problem just by choosing a relative small number β_{l_0}. A statistical justification of a design performing well for the estimation of completely independent parameters in different models can be given under special circumstances. As an example we will consider the following two regression setups. Assume that an estimate of the (unknown) regression function $g(x) \in \mathcal{F}_m$ at a given point $x_0 \notin x$ is desired. In this case we choose $c_l = f_l(x_0)$ $(l = 1, \ldots, m)$ and obtain an extrapolation design robust with respect to the model assumptions. If the experimenter has a very precise knowledge about the underlying model and wants to estimate the regression function at the points $x_1, \ldots, x_m$ we choose $c_l = f_1(x_l)$ $(l = 1, \ldots, m)$ and all models g_l equal g_1. Secondly, consider the case of nested regression models (i.e. $f_l(x) = (f_{11}(x), \ldots, f_{1l}(x))'$, $l = 1, \ldots, m$) and define $c_l = (0, \ldots, 0, 1)' \in \mathbb{R}^l$. In this case an $\Phi^c_{p,\beta}$-optimal design may be useful to decide how many regression functions have to be included into the model.

For the classical c_1-optimality criterion (i.e. $m = 1$ and $\beta_1 = 1$ in Definition 1.1) the theorem of Elfving (1952) is an important tool for the (geometric) characterization of the c_1-optimal design (see also Karlin and Studden (1966) and Pukelsheim (1981) for more details). A c_1-optimal design can be characterized as a design which allows a representation of the intersection of the half line $\{\lambda c_1 \mid \lambda > 0\}$ with the boundary of the Elfving set

$$\mathscr{R}_1 = \text{co}\,(\{f_1(x) \mid x \in \mathscr{X}\} \cup \{-f_1(x) \mid x \in \mathscr{X}\}) \subseteq \mathbb{R}^{k_1} \tag{1.2}$$

where co $(\mathscr{A})$ denotes the convex hull of a set $\mathscr{A} \subseteq \mathbb{R}^{k_1}$. In Section 2 we will present a generalization of Elfving's Theorem for the $\Phi^c_{p,\beta}$-optimal design problem of Definition 1.1. In the same section the relationship between $\Phi^c_{p,\beta}$-optimal designs with respect to different values of $p \in (-\infty, 1]$ is given and a sufficient condition is stated when a design is $\Phi^c_{p,\beta}$-optimal for all $p \in [-\infty, 1]$. Section 3 deals with the E-optimality criterion and its relation to the $\Phi^c_{p,\beta}$-optimal design problem where all functions f_l are identical (i.e. $f_1(x) = f_l(x)$, $l = 1, \ldots, m$). We show that the E-optimal design is $\Phi^c_{p,\beta}$-optimal with respect to a prior β, where $p \in (-\infty, 1]$ is arbitrary, $c = (c_1', \ldots, c_{k_1}')'$ is any inball vector of the generalized Elfving set and the l-th weight of the prior β is proportional to $(c_l' c_l)^{p+1}$ $(l = 1, \ldots, k_1)$.

2 $\Phi^c_{p,\beta}$-Optimal Designs

In the following we will define a generalized Elfving set $\mathscr{R}^{\beta}_m$ corresponding to the $\Phi^c_{p,\beta}$-optimal design problem by

$$\mathscr{R}^{\beta}_m := \text{co}\left(\left\{(\varepsilon_1 f_1'(x), \ldots, \varepsilon_m f_m'(x))' \mid x \in \mathscr{X},\ \varepsilon_l \in \left[-\frac{1}{\sqrt{\beta_l}}, \frac{1}{\sqrt{\beta_l}}\right],\ \sum_{l=1}^{m} \beta_l \varepsilon_l^2 = 1\right\}\right) \tag{2.1}$$

which is a convex, symmetric and compact subset of $\mathbb{R}^k$ $\left(k = \sum_{l=1}^{m} k_l\right)$ containing the point 0. Note that in the case $m = 1$ $(\beta_1 = 1)$ (2.1) gives exactly the set defined in (1.2). In general the structure of $\mathscr{R}^{\beta}_m$ is very complicated and is illustrated by some examples in a paper of Dette (1992). We are now able to state an analogous geometric characterization of the $\Phi^c_{p,\beta}$-optimal design problem for

the class $\mathcal{F}_m$ with respect to a prior β as given by Elfving (1952). The proof of the following theorem is performed by similar arguments as given in Dette (1992) (where the case $p = 0$ is considered) and therefore omitted.

Theorem 2.1: Let $p \in (-\infty, 1]$, *a design* $\xi = \begin{Bmatrix} x_v \\ p_v \end{Bmatrix}_{v=1}^{s}$ *(for which* $c_l'\theta_l$ *is estimable* $l = 1, \ldots, m$*) is* $\Phi_{p,\beta}^{c}$*-optimal for the class* $\mathcal{F}_m$ *with respect to the prior* β *if and only if there exist positive numbers* $\gamma_1, \ldots, \gamma_m$ *and numbers* $\varepsilon_{11}, \ldots, \varepsilon_{1s}, \varepsilon_{21}, \ldots, \varepsilon_{2s}, \ldots, \varepsilon_{m1}, \ldots, \varepsilon_{ms}$ *such that the following properties* (A), (B), (C) *and* (D) *hold.*

$$\gamma_l c_l = \sum_{v=1}^{s} p_v \varepsilon_{lv} f_l(x_v) \qquad l = 1, \ldots, m \ . \tag{A}$$

The point $(\gamma_1 c_1', \ldots, \gamma_m c_m')'$ *is a boundary point of the set* $\mathcal{R}_m^{\beta}$ *with a supporting hyperplane* $(\beta_1 d_1', \ldots, \beta_m d_m')'$. (B)

$$\gamma_l^2 \left(\frac{c_l' d_l}{\gamma_l} \right)^{p+1} = \left[\sum_{j=1}^{m} \beta_j \left(\frac{\gamma_j}{c_j' d_j} \right)^{p} \right]^{-1} \qquad l = 1, \ldots, m \tag{C}$$

$$\sum_{l=1}^{m} \beta_l \varepsilon_{lv}^2 = 1 \qquad v = 1, \ldots, s \ . \tag{D}$$

Remark 2.2: It follows from the proof of Theorem 2.1 that the quantities γ_l, ε_{lv} and d_l are given by

$$\gamma_l^{-2} = (c_l' M_l^{-}(\xi) c_l)^{p+1} \cdot \sum_{j=1}^{m} \beta_j (c_j' M_j^{-}(\xi) c_j)^{-p} \qquad (l = 1, \ldots, m) \tag{2.2}$$

$$d_l = \gamma_l G_l c_l \qquad (l = 1, \ldots, m) \tag{2.3}$$

$$\varepsilon_{lv} = f_l'(x_v) d_l \qquad (l = 1, \ldots, m,\ v = 1, \ldots, s)$$

where the matrices G_l are "suitable" generalized inverses of $M_l(\xi)$, $l = 1, \ldots, m$ (see Dette (1992) for more details). Note also that the numbers $\gamma_1, \ldots, \gamma_m$ are uniquely determined which is a consequence of (2.2) and the concavity of the function $\Phi_{p,\beta}^{c}(\xi)$.

Remark 2.3: The case $p = -1$ plays a particular role in Theorem 2.1 because in this case all quantities γ_l in the general Elfving Theorem coincide, that is

$$\gamma^{-2} = \gamma_l^{-2} = \sum_{j=1}^{m} \beta_j c_j' M_j^- (\xi) c_j \quad l = 1, \ldots, m \ .$$

Here the $\Phi^c_{-1,\beta}$-optimal design can be determined (as in the classical case $m = 1$) finding a representation (A) of the point where the line $\{\lambda (c_1', \ldots, c_m')' \mid \lambda > 0\}$ intersects the boundary of the Elfving set $\mathcal{R}_m^\beta$. For the choice of identical models we thus get the Elfving theorem for $A'\theta$-optimality given in Studden (1971) where $A = A_\beta = (\sqrt{\beta_1} c_1, \ldots, \sqrt{\beta_m} c_m)$.

Theorem 2.4: Let $p_1 \in (-\infty, 1]$ and ξ denote a $\Phi^c_{p_1,\beta}$-optimal design for the class $\mathcal{F}_m$ with respect to the prior $\beta = (\beta_1, \ldots, \beta_m)$ where $\gamma_1, \ldots, \gamma_m$, $d_1, \ldots, d_m$ are the quantities of the Elfving Theorem 2.1. For every $p_2 \in (-\infty, 1]$ the design ξ is also $\Phi^c_{p_2,\tilde{\beta}}$-optimal for the class $\mathcal{F}_m$ with respect to the prior $\tilde{\beta} = (\tilde{\beta}_1, \ldots, \tilde{\beta}_m)$ where

$$\tilde{\beta}_l = \beta_l \frac{\gamma_l^{p_1 - p_2} (c_l' d_l)^{p_2 - p_1}}{\sum_{l=1}^{m} \beta_l \gamma_l^{p_1 - p_2} (c_l' d_l)^{p_2 - p_1}} \quad (l = 1, \ldots, m) \ .$$

If $p \in (-\infty, 1]$ and ξ denotes a $\Phi^c_{p,\beta}$-optimal design for the class $\mathcal{F}_m$ with respect to the prior β such that the quantities $c_l' d_l / \gamma_l$ are constant, then the design ξ is $\Phi^c_{q,\beta}$-optimal for the class $\mathcal{F}_m$ with respect to the prior β whenever $q \in (-\infty, 1]$. If, additionally, the matrices $M_1(\xi), \ldots, M_m(\xi)$ are nonsingular the design ξ is also $\Phi^c_{-\infty,\beta}$-optimal for the class $\mathcal{F}_m$.

Proof: Let $k = \sum_{l=1}^{m} k_l$ and

$$\tilde{M}(\xi) = \begin{pmatrix} M_1(\xi) & & \\ & \ddots & \\ & & M_m(\xi) \end{pmatrix} \in \mathbb{R}^{k \times k} \ ,$$

$$\tilde{K} = \begin{pmatrix} \tilde{c}_1 & & \\ & \ddots & \\ & & \tilde{c}_m' \end{pmatrix} \in \mathbb{R}^{m \times k} \ ,$$

(all other entries in the matrices are 0) $\tilde{c}_l = \beta_l^{-1/2p} c_l$ $(l = 1, \ldots, m)$ and $C_{\tilde{K}}(\tilde{M}) = \min\{L \tilde{M} L' \mid L \tilde{K}' = I_m\}$, then minimizing $\Phi^c_{p,\beta}$ over all designs ξ is equivalent to maximizing the (unweighted) p-mean $(\text{trace}\, (C_{\tilde{K}}(\tilde{M})^p)^{1/p}$ in the convex compact set of associated matrices $\tilde{M}(\xi)$. Thus the first part of the theorem follows im-

mediately by computing gradients of $\Phi^c_{p,\beta}$ via standard rules of differentiation (see e.g. Gaffke (1985, 1987)). For the second part we use (2.3) and obtain by an application of the first part ($p_1 = p$, $p_2 = q$) that the design ξ is $\Phi^c_{q,\beta}$-optimal for the class $\mathcal{F}_m$ with respect to the prior β whenever $q \in (-\infty, 1]$. Finally, the case $q = -\infty$ follows considering the limit $q \to -\infty$. ■

We will finish this section showing that *every* design which allows a representation of a boundary point $\gamma c = (\gamma_1 c_1', \ldots, \gamma_m c_m')'$ of the generalized Elfving set $\mathcal{R}^\beta_m$ is already $\Phi^c_{p,\beta}$-optimal for a special prior β.

Theorem 2.5: Let $c = (c_1', \ldots, c_m')'$ *and* $\gamma_c = (\gamma_1 c_1', \ldots, \gamma_m c_m')'$ *denote an arbitrary boundary point of the set* $\mathcal{R}^\beta_m$ *with supporting hyperplane* $(\beta_1 d_1', \ldots \beta_m d_m')'$. *Let* $\xi = \begin{Bmatrix} x_\nu \\ p_\nu \end{Bmatrix}_{\nu=1}^{s}$ *denote a design which allows a representation*

$$\gamma_l c_l = \sum_{\nu=1}^{s} p_\nu \varepsilon_{l\nu} f_l(x_\nu) \qquad (l = 1, \ldots, m) \ ,$$

such that for $\nu = 1, \ldots, s$ $\sum_{l=1}^{m} \beta_l \varepsilon_{l\nu}^2 = 1$. *For every* $p \in (-\infty, 1]$ *the design* ξ *is* $\Phi^c_{p,\beta}$*-optimal for the class* $\mathcal{F}_m$ *with respect to the prior* $\tilde{\beta} = (\tilde{\beta}_1, \ldots, \tilde{\beta}_m)$ *where*

$$\tilde{\beta}_l = \beta_l \gamma_l^2 \left(\frac{c_l' d_l}{\gamma_l} \right)^{p+1} \left[\sum_{j=1}^{m} \beta_j \gamma_j^2 \left(\frac{c_j' d_j}{\gamma_j} \right)^{p+1} \right]^{-1} \qquad (l = 1, \ldots, m) \ .$$

Proof: By an application of Theorem 2.1 in the case $p = -1$ we obtain that the design ξ is $\phi^{\tilde{c}}_{-1,\beta}$-optimal for the class $\mathcal{F}_m$ with respect to the prior β where $\tilde{c} = (\gamma_1 c_1', \ldots, \gamma_m c_m')'$ (note that in the case $p = -1$ the condition (C) of Theorem 2.1 is evident by the property (B)). The assertion now follows directly from Theorem 2.4 putting $p_1 = -1$ and $p_2 = p$. ■

3 E-Optimal Designs

In this section we will investigate the relation between the E-optimality criterion and the model robust criterion given in Definition 1.1. In what follows $\lambda_{\min}$ always denote the minimum eigenvalue of the information matrix corresponding

to the E-optimal design and $\lambda_{\min}(A)$ denotes the minimum eigenvalue of a symmetric matrix $A \in \mathbb{R}^{k_1 \times k_1}$. Pukelsheim (1980) showed that ξ is E-optimal iff there exists a matrix

$$E = \sum_{i=1}^{k_0} \alpha_i z_i z_i' \in \mathbb{R}^{k_1 \times k_1} \tag{3.1}$$

such that $f_1'(x) E f_1(x) \leq \lambda_{\min}(M_1(\xi))$ for all $x \in \mathcal{X}$. The vectors z_i in the representation (3.1) are normalized eigenvectors ($||z_i||_2 = 1$) corresponding to the minimum eigenvalue $\lambda_{\min}$ of $M_1(\xi)$, the α_i are positive numbers with sum 1 and $k_0 = \text{rank}(E) \leq k_1$. Throughout this section we will consider the model robust optimality criterion of Definition 1.1 where all models $g_l \in \mathcal{F}_m$ are identical (that is $f_l(x) = f_1(x)$, $l = 1, \ldots, m$). The following result generalizes Theorem 3.2 of Dette and Studden (1992) ($p = -1$).

Theorem 3.1: For every $p \in [-\infty, 1]$ the design ξ is $\Phi^c_{p,\beta}$-optimal with respect to the prior α where $\alpha = (\alpha_1, \ldots, \alpha_{k_0})$ and the vector $c = (z_1', \ldots, z_{k_0}')'$ are defined by (3.1).

For the following investigations of the geometric properties of the E-optimal design we will consider the corresponding Elfving set $\mathcal{R}^{\alpha}_{k_0}$ defined in (2.1) and its "inball radius"

$$r_{k_0} := \min\{||c||_2 \mid c \in \partial \mathcal{R}^{\alpha}_{k_0}\} . \tag{3.2}$$

Because $\mathcal{R}^{\alpha}_{k_0}$ is a convex subset of $\mathbb{R}^{k_0 k_1}$ we obtain an equivalent representation of (3.2) by the distances of covering halfspaces to $\mathcal{R}^{\alpha}_{k_0}$

$$r_{k_0} = \min\left\{ \frac{1}{\sqrt{\sum_{l=1}^{k_0} d_l' d_l}} \,\middle|\, d_l \in \mathbb{R}^{k_1}, \left| \sum_{l=1}^{k_0} \varepsilon_l d_l' f_1(x) \right| \leq 1, \ \forall x \in \mathcal{X}, \ \sum_{l=1}^{k_0} \alpha_l \varepsilon_l^2 = 1 \right\}. \tag{3.3}$$

Every vector $c \in \partial \mathcal{R}^{\alpha}_{k_0}$ with $||c||_2 = r_{k_0}$ is called an "inball vector" of $\mathcal{R}^{\alpha}_{k_0}$.

Theorem 3.2: Let E be defined by (3.1) then the point $\sqrt{\lambda_{\min}}(z_1', \ldots, z_{k_0}')'$ is a boundary point of the set $\mathcal{R}^{\alpha}_{k_0}$ with supporting hyperplane $\frac{1}{\sqrt{\lambda_{\min}}}(\alpha_1 z_1', \ldots, \alpha_{k_0} z_{k_0}')'$. For the inball radius r_{k_0} of $\mathcal{R}^{\alpha}_{k_0}$ we have the estimates

$$\lambda_{\min} \leq r_{k_0}^2 \leq k_0 \lambda_{\min} \ . \tag{3.4}$$

If, additionally, all weights α_l in (3.1) are equal the vector $c = \sqrt{\lambda_{\min}}\,(z_1', \ldots, z_{k_0}')'$ defines an inball vector of $\mathcal{R}_{k_0}^{\alpha^*}\left(\alpha^* = \left(\frac{1}{k_0}, \ldots, \frac{1}{k_0}\right)\right)$ and the inball radius is given by $r_{k_0} = k_0 \lambda_{\min}$.

Proof: Consider the bijective linear mapping

$$\mathrm{vec}_\alpha(A) = \left(\frac{1}{\sqrt{\alpha_1}} a_1', \ldots, \frac{1}{\sqrt{\alpha_{k_0}}} a_{k_0}'\right)' , \quad A = (a_1, \ldots, a_{k_0}) \in \mathbb{R}^{k_1 \times k_0}$$

from $\mathbb{R}^{k_1 \times k_0}$ onto $\mathbb{R}^{k_1 k_0}$. It is easy to see that

$$\mathrm{vec}_\alpha(S_{k_0}) = \mathcal{R}_{k_0}^{\alpha} , \quad \mathrm{vec}_\alpha(\partial S_{k_0}) = \partial \mathcal{R}_{k_0}^{\alpha} \tag{3.5}$$

where $S_{k_0} := \mathrm{co}\{f_1(x)\varepsilon' \mid x \in \mathcal{X},\ \varepsilon \in \mathbb{R}^{k_0},\ \|\varepsilon\|_2 = 1\} \subseteq \mathbb{R}^{k_1 \times k_0}$ denotes the Elfving set investigated by Studden (1971). By straight forward calculations it follows that $A = (\sqrt{\alpha_1} a_1, \ldots, \sqrt{\alpha_{k_0}} a_{k_0})$ is a covering halfspace to S_{k_0} if and only if $a = (\alpha_1 a_1', \ldots, \alpha_{k_0} a_{k_0}')'$ is a covering halfspace to $\mathcal{R}_{k_0}^{\alpha}$. Dette and Studden (1992) showed that the squared inball radius of S_{k_0} is given by $\lambda_{\min}$. Thus we obtain from

$$\lambda_{\min} \leq \frac{1}{\mathrm{trace}\,(A'A)} = \frac{1}{\sum_{l=1}^{k_0} \alpha_l a_l' a_l} \leq \frac{1}{a'a}$$

and from (3.3) the inequality $\lambda_{\min} \leq r_{k_0}^2$. It also follows from Theorem 3.3 of Dette and Studden (1992) that the point $Z = \sqrt{\lambda_{\min}}\,(\sqrt{\alpha_1} z_1, \ldots, \sqrt{\alpha_{k_0}} z_{k_0})$ defines an inball vector of the set S_{k_0} (here z_l, α_l and k_0 are the quantities of the representation of the matrix E in (3.1)). By (3.5) $\mathrm{vec}_\alpha(Z)$ is a boundary point of the set $\mathcal{R}_{k_0}^{\alpha}$ with norm $\|\mathrm{vec}_\alpha(Z)\|_2^2 = k_0 \lambda_{\min}$ which proves the right inequality of (3.4) (note that $\mathrm{vec}_\alpha(Z)$ is not necessarily an inball vector of $\mathcal{R}_{k_0}^{\alpha}$ because the supporting hyperplane $d = \frac{1}{\sqrt{\lambda_{\min}}}\,(\alpha_1 z_1', \ldots, \alpha_{k_0} z_{k_0}')'$ to $R_{k_0}^{\alpha}$ at the point $\mathrm{vec}_\alpha(Z)$ has in general a different direction as the vector $\mathrm{vec}_\alpha(Z)$).

In the case of equal $\alpha_l = \frac{1}{k_0}$ $(l = 1, \ldots, k_0)$ we have $||\text{vec}_\alpha(A)||_2^2 = k_0 ||A||_2^2$ and obtain from (3.5) the second assertion of the theorem. ■

The results stated so far are more of theoretical interest for the determination of the E-optimal design because they assume the knowledge of $k_0, \alpha_1, \ldots, \alpha_{k_0}$ in the representation (3.1) which is usually not available without knowing the E-optimal design. Nevertheless they provide a first insight into the geometric structure of the problem and also indicate the usefulness of a representation with equal weights $\alpha_l = \frac{1}{k_0}$ $(l = 1, \ldots, k_0)$. The following Lemma shows that there always exists such a representation.

Lemma 3.3: For every matrix E in (3.1) there always exists a representation with identical weights α_l.

Proof: Considering a spectral decomposition of the matrix E in (3.1) we obtain a representation with orthonormal vectors z_j $(k_0 = \text{rank}(E) \leq k_1)$. Let $k^* = 2^{k_0 - 1}$ and

$$F = \begin{pmatrix} 1 & 1 & \ldots & 1 & 1 \\ 1 & 1 & \ldots & 1 & -1 \\ 1 & 1 & \ldots & -1 & 1 \\ 1 & 1 & \ldots & -1 & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & -1 & \ldots & 1 & 1 \\ 1 & -1 & \ldots & 1 & -1 \\ 1 & -1 & \ldots & -1 & 1 \\ 1 & -1 & \ldots & -1 & -1 \end{pmatrix} \in \mathbb{R}^{k^* \times k_0}$$

where all $2^{k_0 - 1}$ components of the first column of F are 1, the first $2^{k_0 - 2}$ components of the second column are 1 while the second $2^{k_0 - 2}$ components of this column are -1 and so on.... Defining new vectors $\tilde{z}_j \in \mathbb{R}^{k_1}$ by

$$(\tilde{z}_1, \ldots, \tilde{z}_{k^*}) = (\sqrt{\alpha_1} z_1, \ldots, \sqrt{\alpha_{k_0}} z_{k_0}) F'$$

(e.g. $\tilde{z}_1 = \sum_{i=1}^{k_0} \sqrt{\alpha_i} z_i$, $\tilde{z}_{k^*} = \sqrt{\alpha_1} z_1 - \sum_{i=2}^{k_0} \sqrt{\alpha_i} z_i$) we have

$$\frac{1}{k^*}\sum_{j=1}^{k^*}\tilde{z}_j\tilde{z}_j' = \frac{1}{k^*}(\tilde{z}_1,\ldots,\tilde{z}_{k^*})\begin{pmatrix}\tilde{z}'\\ \vdots\\ \tilde{z}'_{k^*}\end{pmatrix}$$

$$= \frac{1}{k^*}(\sqrt{\alpha_1}z_1,\ldots,\sqrt{\alpha_{k_0}}z_{k_0})F'F\begin{pmatrix}\sqrt{\alpha_1}z_1'\\ \vdots\\ \sqrt{\alpha_{k_0}}z'_{k_0}\end{pmatrix}$$

$$= (\sqrt{\alpha_1}z_1,\ldots,\sqrt{\alpha_{k_0}}z_{k_0})\begin{pmatrix}\sqrt{\alpha_1}z_1'\\ \vdots\\ \sqrt{\alpha_{k_0}}z'_{k_0}\end{pmatrix} = \sum_{i=1}^{k_0}\alpha_i z_i z_i' = E\ .$$

The vectors $\tilde{z}_j$ are eigenvectors of the information matrix of the E-optimal design corresponding to the minimum eigenvalue $\lambda_{\min}$. From the orthonormality of the vectors z_j we obtain $||\tilde{z}_j||_2 = 1$ which proves the assertion of Lemma 3.3. ■

In view of Theorem 3.2 and Lemma 3.3 it now makes sense to investigate the inball radii of the sequence of Elfving sets $\left(\alpha^* = \left(\frac{1}{m},\ldots,\frac{1}{m}\right)\right)$

$$\mathcal{R}_m^{\alpha^*} = \operatorname{co}\left(\left\{(\varepsilon_1 f_1'(x),\ldots,\varepsilon_m f_1'(x))' \mid x\in\mathcal{X},\ \varepsilon_l\in[-\sqrt{m},\sqrt{m}],\ \sum_{l=1}^{m}\varepsilon_l^2 = m\right\}\right)$$

($m = 1,2,\ldots$) for the determination of the minimum eigenvalue of the E-optimal information matrix. By the results stated so far we have equality $r_{k^*}^2 = k^*\lambda_{\min}$ for some $k^* \le 2^{k_1-1}$ (note that $k_0 = \operatorname{rank}(E) \le k_1$). The following theorem shows that this identity holds already for the set $\mathcal{R}_{k_1}^{\alpha^*}$.

Theorem 3.4: Let r_m denote the inball radius of $\mathcal{R}_m^{\alpha^*}$, then we have $r_m^2 = m\lambda_{\min}$ for all $m \ge k_0$. Whenever $m \ge k_0$ and $c = (c_1',\ldots,c_m')'$ is an inball vector of the set $\mathcal{R}_m^{\alpha^*}$ the E-optimal design is $\Phi_{p,\beta}^c$-optimal with respect to the prior $\beta = (\beta_1,\ldots,\beta_m)$ where

$$\beta_l = \frac{(c_l'c_l)^{p+1}}{\sum_{l=1}^{m}(c_l'c_l)^{p+1}} \qquad l = 1,\ldots,m\ . \tag{3.6}$$

Proof: Observing the reasoning in the proof of Theorem 3.2 we see that the Elfving set S_m is a scaled version of the Elfving set $\mathcal{R}_m^{\alpha^*}$ and obtain for the corresponding inball radii

$$s_m = \frac{1}{\sqrt{m}} r_m \quad \text{for all} \quad m \in \mathbb{N} .$$

By Theorem 3.3 of Dette and Studden (1992) we have for $m \geq k_0$ $r_m^2 = m\lambda_{\min}$ which proves the first part of the assertion. For the second part let $m \geq k_0$ and $c = (c_1', \ldots, c_m')'$ denote an inball vector of the set $\mathcal{R}_m^{\alpha *}$ (i.e. $c \in \partial \mathcal{R}_m^{\alpha *}$ and $||c||_2^2 = m\lambda_{\min}$). Using (3.5) it follows that $\tilde{C} = \text{vec}_{\alpha *}^{-1}(c) = \frac{1}{\sqrt{m}}(c_1, \ldots, c_m) \in \mathbb{R}^{k_1 \times m}$ is a boundary point of S_m with norm $||\tilde{C}||_2^2 = \text{trace}\,(\tilde{C}'\tilde{C}) = \frac{1}{m}||c||_2^2 = \lambda_{\min}$. Thus the matrix $\tilde{C}$ is also an inball vector of S_m and we obtain from Theorem 3.4 of Dette and Studden (1992) that the E-optimal design ξ is an optimal design for $\tilde{C}'\theta_1$. This implies that the E-optimal design is also $\Phi_{-1,\beta}^c$-optimal with respect to the prior $\beta^* = \left(\frac{1}{m}, \ldots, \frac{1}{m}\right)$ and the quantities of the Elfving Theorem 2.1 are given by $\gamma_l = 1$ and $d_l = \frac{c_l}{\lambda_{\min}}$ $(l = 1, \ldots, m)$. By an application of Theorem 2.4 it now follows that for every $p \in (-\infty, 1]$ the E-optimal design is also $\Phi_{p,\beta}^c$-optimal with respect to the prior $\beta = (\beta_1, \ldots, \beta_m)$ given in (3.6). ■

Using the above results the E-optimal design can now be determined in the following way. In a first step we have to find an inball vector c and its supporting hyperplane of the set $\mathcal{R}_{k_1}^{\alpha *}$ (note that the number $k_0 = \text{rank}\,(E)$ in the Theorem 3.4 is unkown but always $\leq k_1$). This problem is equivalent to the calculation of the inball vector and its supporting hyperplane of the set S_{k_1} (which is a scaled version of $\mathcal{R}_{k_1}^{\alpha *}$). In a second step we have to find the $\Phi_{p,\beta}^c$-optimal designs with respect to the prior β in (3.6) using the results on $\Phi_{p,\beta}^c$-optimality. By Theorem 3.4 the E-optimal design has to be among these designs. In general the computational difficulties in this part will depend sensitively on the choice of the parameter p in the optimality criterion. Usually the cases $p = 0$ (for $p = 0$ the criterion is invariant with respect to different scalings of the vectors c_l) and $p = -1$ (for $p = -1$ all numbers γ_l in Theorem 2.1 are equal, see Remark 2.3) can be treated easier compared to other choices of p. A couple of examples for this procedure in the case $p = -1$ can be found in Dette and Studden (1992) and we will present here a simple example where the choice $p = 0$ might be useful.

Example 3.5: Consider the weighted polynomial regression model $f(x) = \sqrt{b^2 - x^2} \times (1, x)'$ on the interval $[-b, b]$ where $b > \sqrt{2}$. Observing (3.3) it is straight forward to show that an inball vector of $\mathcal{R}_2^{\alpha *}$ is given by $c = (c_1', c_2')' = \sqrt{2}\,(0, 1, \sqrt{b^2 - 2}, 0)'$. From standard arguments of optimal design theory it follows that

there exists a symmetric E-optimal design. Expressing the information matrix of such a symmetric design in terms of canonical moments of ξ (see Studden (1980), p. 1336) we obtain ($q_2 = 1-p_2$, $q_4 = 1-p_4$)

$$M_2(\xi) = \begin{pmatrix} b^2 q_2 & 0 \\ 0 & b^4 q_2 p_2 q_4 \end{pmatrix}.$$

Therefore a $\Phi^c_{0,\hat{\beta}}$-optimal design with respect to the prior $\hat{\beta} = (1/(b^2-1), (b^2-2)/(b^2-1))$ maximizes the product $q_2(p_2 q_4)^{1/(b^2-1)}$ and we obtain for the canonical moments of the maximizing measure $p_2 = 1/b^2$, $p_4 = 0$. Using the results of Studden (1982) (transformed to the interval $[-b, b]$) we obtain that the unique $\Phi^c_{0,\hat{\beta}}$-optimal design puts equal masses at the points -1 and 1 (independent of the length of the interval $[-b, b]$). By Theorem 3.4 this design must be E-optimal too. ■

If the E-optimal design and the matrix E in (3.1) are known it follows from Theorem 3.1 that the E-optimal design also maximizes every (weighted) p-mean of the inverse variances of the least squares estimates for the linear combinations $z_l'\theta_1$ ($l = 1, \ldots, k_0$) and we obtain a new interpretation of optimality for the E-optimal design. It is remarkable that by Theorem 3.4 the E-optimal design is not necessarily $\Phi^c_{p,\beta}$-optimal with respect to the uniform prior $\alpha^* = \left(\frac{1}{m}, \ldots, \frac{1}{m}\right)$ (although c is an inball vector of the set $\mathcal{R}^{\alpha^*}_m$ ($m \geq k_0$)). The weights of the prior in the $\Phi^c_{p,\beta}$-optimality criterion will usually depend on the length of the components c_l of the inball vector $c = (c_1', \ldots, c_m')'$ and these are not necessarily all equal (see Example 3.5).

Acknowledgements: This paper is based on a part of the author's Habilitationsschrift written at the Universität Göttingen. The author would like to thank an anonymous referee for his careful reading of the manuscript and his helpful comments.

References

Dette H (1992) Elfving's Theorem for D-Optimality. Ann Statist, to appear (Technical Report # 16-91, Department of Statistics, Purdue University)

Dette H, Studden WJ (1992) Geometry of E-Optimality. Ann Statist, to appear (Technical Report # 31-91, Department of Statistics, Purdue University)

Elfving G (1952) Optimum Allocation in Linear Regression. Ann Math Statist 23:255–262

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 Admissibility for Partial Parameter Estimation in Linear Regression Models. Ann Statist 15:942–957

Karlin S, Studden WJ (1966) Optimal Experimental Designs. Ann Math Statist 37:783–815

Krafft O (1978) Lineare statistische Modelle und optimale Versuchspläne. Vandenhoeck & Ruprecht, Göttingen

Pukelsheim F (1980) On Linear Regression Designs which Maximize Information. J Statist Planning and Inference 4:339–364

Pukelsheim F (1981) On c-Optimal Design Measures. Math Operationsforsch. Statist Ser Statist 12:13–20

Studden WJ (1971) Elfving's Theorem and Optimal Designs for Quadratic Loss. Ann Math Statist 42:1613–1621

Studden WJ (1980) D_s-Optimal Designs for Polynomial Regression Using Continued Fractions. Ann Statist 8:1132–1141

Studden WJ (1982) Optimal Desings for Weighted Polynomial Regression Using Canonical Moments. In: Statistical Decision Theory and Related Topics III. Academic Press, New York, pp 335–350

Received 16 December 1991
Revised version 18 May 1992