Metrika (2001) 53: 207–222

Metrika

A stochastic characterization of Loewner optimality design criterion in linear models

Erkki P. Liski[1], **Alexander Zaigraev**[2]

[1] Department of Mathematics, Statistics and Philosophy, University of Tampere, P.O. BOX 607, FIN-33101 Tampere, Finland (e-mail: epl@uta.fi)
[2] Faculty of Mathematics and Computer Science, Nicholas Copernicus University, Chopin str. 12/18, PL-87-100 Torun, Poland

Received: 9 August 2000

Abstract. In this paper we present a new stochastic characterization of the Loewner optimality design criterion. The result is obtained by proving a generalization to the well known corollary of Anderson's theorem. Certain connections between the Loewner optimality and the stochastic distance optimality design criterion are showed. We also present applications and generalizations of the main result.

Key words: Anderson's theorem, stochastic convex and distance optimality criteria, Kiefer optimality, admissibility, information equivalence, multifactor first degree polynomial models, orthogonal and simplex designs

1 Introduction

There exists an extensive literature on optimal design criteria. For references see Shah and Sinha (1989) and Pukelsheim (1993), for example. Among them there are traditional criteria like A-, D- or E-optimality and more sofisticated ones like Kiefer optimality or Loewner optimality which is quite a strong criterion. However, there are models, as e.g. two-way classification models (see Pukelsheim 1993, Section 4.8), for which a Loewner optimal design exists.

In this paper, we assume the classical linear model

$$\mathbf{Y} \sim \mathrm{N}_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}_n), \tag{1}$$

where the $n \times 1$ response vector $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)'$ follows a multivariate normal distribution, $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)'$ is the $n \times k$ model matrix of the full rank k, $k \leq n$, $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_k)'$ is the $k \times 1$ parameter vector, $\mathrm{E}(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$ is the expectation vector of $\mathbf{Y}$ and $\mathrm{D}(\mathbf{Y}) = \sigma^2\mathbf{I}_n$ is the dispersion matrix of $\mathbf{Y}$, where $\sigma^2 = \mathrm{V}(Y_i)$ for every $i = 1, 2, \ldots, n$ and $\mathbf{I}_n$ is the $n \times n$ identity matrix.

An experimental design $\xi_{(n)}$ for a given number n of trials specifies $l \leq n$ distinct regression vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_l$ and assigns to them frequences n_i such that $\sum_{i=1}^{l} n_i = n$. The regression vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_l$ are called the support of the design $\xi_{(n)}$. A design assigns the weight $\frac{n_i}{n}$ to each vector $\mathbf{x}_i$, $i = 1, 2, \ldots, l$. Such designs are called *exact*.

Let $\hat{\boldsymbol{\beta}}$ be the least squares estimator (LSE) of $\boldsymbol{\beta}$ which is the best linear unbiased estimator (BLUE) of $\boldsymbol{\beta}$. The dispersion matrix of $\hat{\boldsymbol{\beta}}$ is $\mathrm{D}(\hat{\boldsymbol{\beta}}) = \frac{\sigma^2}{n}\mathbf{M}^{-1}$, where the matrix $\mathbf{M} = \sum_{i=1}^{l} \frac{n_i}{n}\mathbf{x}_i\mathbf{x}_i'$ is the *moment matrix* of $\xi_{(n)}$.

More generally, we may allow the weights vary continuously in the interval $[0, 1]$. In this case we deal with designs for a infinite number of trials (see e.g. Pukelsheim 1993, Section 1.24). Such designs are called *continuous*. Each continuous design ξ is a discrete probability measure taking values $p_i \geq 0$ at vectors $\mathbf{x}_i$, $i = 1, 2, \ldots, l$, that is

$$\xi = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_l; p_1, p_2, \ldots, p_l\}, \quad \sum_{i=1}^{l} p_i = 1.$$

The moment matrix of a design ξ is also defined by $\mathbf{M}(\xi) = \sum_{i=1}^{l} p_i\mathbf{x}_i\mathbf{x}_i'$.

The problem of searching for an optimal design is simplified by considering continuous designs for an infinite number of trials, thus ignoring the constraint that the number of trials at any design point must be an integer. In practice all designs are exact. However, continuous design can be used to approximate an exact design.

2 Loewner dominance

For describing the notion of Loewner dominance, let us first consider a line fit model. Suppose we have $n \geq 2$ uncorrelated responses

$$Y_{ij} = \beta_1 + \beta_2 x_i + E_{ij}, \quad i = 1, 2, \ldots, l;\ j = 1, 2, \ldots, n_i \tag{2}$$

with expectations and variances $E(Y_{ij}) = \beta_1 + \beta_2 x_i$ and $V(Y_{ij}) = \sigma^2$, respectively.

An experimental design ξ specifies distinct values $x_1, x_2, \ldots, x_l$ chosen from a given experimental domain (usually an interval $[a, b]$) and assigns to them weights p_i such that $\sum_{i=1}^{l} p_i = 1$. The moment matrix of a design ξ is given as

$$\mathbf{M}(\xi) = \sum_{i=1}^{l} p_i\mathbf{x}_i\mathbf{x}_i' = \begin{pmatrix} 1 & \sum_{i=1}^{l} p_i x_i \\ \sum_{i=1}^{l} p_i x_i & \sum_{i=1}^{l} p_i x_i^2 \end{pmatrix},$$

where $\mathbf{x}_i = (1, x_i)'$.

Let $\xi_{x,y;p}$ denotes a 2-point design $\{x, y; p, 1-p\}$ for the LSE of $\boldsymbol{\beta}$ in (2) with weights $0 < p < 1$ and $1-p$ at the points x and y, respectively, $a \le x < y \le b$. By De la Garza (1954), for any l-point design ξ for the LSE of $\boldsymbol{\beta}$ in (2), there exists a 2-point design $\xi_{x,y;p}$ such that $\min\{x_1, x_2, \ldots, x_l\} \le x < y \le \max\{x_1, x_2, \ldots, x_l\}$ and $\mathbf{M}(\xi_{x,y;p}) = \mathbf{M}(\xi)$.

Designs having equal moment matrices are called *information equivalent*. Since comparison between designs in this paper will be based solely on their moment matrices, we can confine our study of a line fit model (2) with the class of 2-point designs.

We say that a design ξ_1 *dominates* a design ξ_2 *in the Loewner ordering sense* if $\mathbf{M}_1 - \mathbf{M}_2$ is a nonnegative definite matrix, where $\mathbf{M}_i$, $i = 1, 2$, are the moment matrices of the designs ξ_1 and ξ_2, respectively. We also denote $\mathbf{M}_1 \ge \mathbf{M}_2$ or $\mathbf{M}_1 - \mathbf{M}_2 \ge \mathbf{0}$ when $\mathbf{M}_1 - \mathbf{M}_2$ is *nonnegative definite* and $\mathbf{M}_1 > \mathbf{M}_2$ or $\mathbf{M}_1 - \mathbf{M}_2 > \mathbf{0}$ when $\mathbf{M}_1 - \mathbf{M}_2$ is *positive definite*. Thus the Loewner partial ordering among moment matrices induces a partial ordering among associated designs. We denote $\xi_1 \succ_L \xi_2$ when ξ_1 dominates ξ_2 with respect to Loewner ordering. If $\xi^* \succ_L \xi$ for all ξ, then ξ^* is *Loewner optimal*. For more extensive discussion on this concept we refer to Marshall and Olkin (1979, p. 462) and Pukelsheim (1993, p. 12 and Chapter 4). The next result shows that there exists a Loewner superior subclass among the 2-point designs.

Lemma 1. *For any given design $\xi_{x,y;r}$, $a \le x < y \le b$ and $a < x$ or $y < b$, for the LSE of $\boldsymbol{\beta}$ in (2), there exists a 2-point design $\xi_p = \{a, b; p, 1-p\}$, $\xi_p \ne \xi_{x,y;r}$, that dominates $\xi_{x,y;r}$, i.e.*

$$\xi_p \succ_L \xi_{x,y;r}.$$

Proof. We have

$$\mathbf{M}(\xi_p) - \mathbf{M}(\xi_{x,y;r}) = \begin{pmatrix} 0 & [pa + (1-p)b] - [rx + (1-r)y] \\ [pa + (1-p)b] - [rx + (1-r)y] & [pa^2 + (1-p)b^2] - [rx^2 + (1-r)y^2] \end{pmatrix}. \tag{3}$$

For any given $\xi_{x,y;r}$ with $a \le x < y \le b$ $(a < x$ or $y < b)$ we can always choose

$$0 < p = \frac{b - [rx + (1-r)y]}{b-a} < 1$$

so that the nondiagonal elements in (3) become zero. Then

$$\begin{aligned} &[pa^2 + (1-p)b^2] - [rx^2 + (1-r)y^2] \\ &\quad = b^2 - [rx^2 + (1-r)y^2] - (a+b)(b - [rx + (1-r)y]) \\ &\quad = r(a+b-x)x + (1-r)(a+b-y)y - ab > 0 \end{aligned}$$

since $(a+b-x)x > ab$ when $x \in (a,b)$. Consequently $\mathbf{M}(\xi_p) - \mathbf{M}(\xi_{x,y;r})$ is nonnegative definite. Thus there always exists $\xi_p \neq \xi_{x,y;r}$ such that $\xi_p \succ_L \xi_{x,y;r}$. $\square$

It is easy to see by (3) that given a design ξ_r, there is no design ξ_p that dominates ξ_r. Thus any 2-point design $\xi_p = \{a, b; p, 1-p\}$ with $0 < p < 1$ is *admissible*. We also say that the moment matrices of those designs are admissible. Correspondingly, each design $\xi_{x,y;r}$ with support points $a \le x < y \le b$ ($a < x$ or $y < b$) is *inadmissible*, and there exists by Lemma 1 an admissible design which dominates it. This means that the admissible designs form a *complete class* (cf. Pukelsheim 1993, Chapter 10).

Sinha (1970) introduced the concept of distance optimality criterion in certain treatment-connected design settings.

Definition 1. Let $\hat{\boldsymbol{\beta}}_1 = \hat{\boldsymbol{\beta}}(\xi_1)$ and $\hat{\boldsymbol{\beta}}_2 = \hat{\boldsymbol{\beta}}(\xi_2)$ be the LSE's of $\boldsymbol{\beta}$ in (1) under the designs ξ_1 and ξ_2, respectively, and $\|\cdot\|$ denotes the Euclidean norm in $\mathbf{R}^k$. If for a given $\varepsilon > 0$

$$\mathrm{P}(\|\hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}\| \le \varepsilon) \ge \mathrm{P}(\|\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}\| \le \varepsilon), \tag{4}$$

then the design ξ_1 is at least as good as ξ_2 with respect to the DS(ε)-criterion.

A design ξ^* is DS(ε)-optimal for the LSE of $\boldsymbol{\beta}$ in the model (1) if it maximizes the probability $\mathrm{P}(\|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\| \le \varepsilon)$. When ξ^* is DS(ε)-optimal for all $\varepsilon > 0$, we say that ξ^* is DS-optimal.

In fact, the relation (4) defines a partial ordering on the set of all possible designs. This kind of partial ordering was referred to as ‘stochastic domination’ in Hwang (1985) or ‘stochastic precision’ in Stępniak (1989). Giovagnoli and Wynn (1995) considered a closely related concept of ordering which they called ‘D-ordering’. In all those papers various aspects of this stochastic ordering were studied mostly from the viewpoint of estimation of linear functions of $\boldsymbol{\beta}$. For completeness one should also mention the review paper by Stępniak and Otachel (1994), and the book by Torgersen (1991, Chapter 8) where the results on comparison of linear experiments were widely discussed. At last, Liski et al. (1999) studied the properties of distance optimality criterion under the classical linear model when observations are independent, homoscedastic and normally distributed.

Further on it is useful to define the DS(ε)-criterion function as

$$\psi_\varepsilon[\mathbf{M}(\xi)] = \mathrm{P}(\|\hat{\boldsymbol{\beta}}(\xi) - \boldsymbol{\beta}\| \le \varepsilon).$$

The DS-criterion is isotonic relative to Loewner ordering (Liski et al. 1999), that is

$$\mathbf{M}(\xi_1) \ge \mathbf{M}(\xi_2) \Rightarrow \psi_\varepsilon[\mathbf{M}(\xi_1)] \ge \psi_\varepsilon[\mathbf{M}(\xi_2)] \quad \text{for all } \varepsilon. \tag{5}$$

Since $\hat{\boldsymbol{\beta}} \sim N_k\left(\boldsymbol{\beta}, \frac{\sigma^2}{n}\mathbf{M}(\xi)^{-1}\right)$ under the model (1) and matrix inversion is antitonic in the case of positive definite matrices, i.e.

$$\mathbf{M}(\xi_1) \geq \mathbf{M}(\xi_2) \Leftrightarrow \mathbf{M}(\xi_1)^{-1} \leq \mathbf{M}(\xi_2)^{-1},$$

the result (5) is a direct consequence of a well-known corollary (see e.g. Perlman 1989, or Tong 1990, Theorem 4.2.5) from Anderson's theorem on the integral of a symmetric unimodal function over a symmetric convex set (see Anderson 1955). We formulate this corollary as Theorem 1.

In the sequel we denote $\mathbf{X} \sim N_k(\mathbf{0}, \mathbf{\Sigma})$ when a $k \times 1$ random vector $\mathbf{X}$ follows a normal distribution with expectation $\mathrm{E}(\mathbf{X}) = \mathbf{0}$ and dispersion matrix $\mathrm{D}(\mathbf{X}) = \mathbf{\Sigma} \geq \mathbf{0}$.

Theorem 1. *Let $\mathbf{X}_1 \sim N_k(\mathbf{0}, \mathbf{\Sigma}_1)$ and $\mathbf{X}_2 \sim N_k(\mathbf{0}, \mathbf{\Sigma}_2)$ be $k \times 1$ normally distributed random vectors, $k \geq 1$, where $\mathbf{\Sigma}_1 > \mathbf{0}$. If $\mathbf{\Sigma}_1 \leq \mathbf{\Sigma}_2$, then*

$$\mathrm{P}(\mathbf{X}_1 \in A) \geq \mathrm{P}(\mathbf{X}_2 \in A)$$

for all convex and symmetric (with respect to the origin) sets $A \subset \mathbf{R}^k$.

In view of equation (3), there is no Loewner optimal design for the LSE of $\boldsymbol{\beta}$ in (2). This result agrees with the more general one from Pukelsheim (1993, Section 4.7).

Moreover, Lemma 1 and relation (5) imply that DS(ε)- and DS-optimal designs are among 2-point designs $\xi_p = \{a, b; p, 1-p\}$ if they exist. We know, in particular, that if $[a, b] = [0, 1]$ there is no DS-optimal design but if $[a, b] = [-1, 1]$ there exists a unique DS-optimal design $\xi_{1/2} = \{-1, 1; p, 1-p\}$ (cf. Liski et al. 1998 and Liski et al. 1999).

In the next section we prove the converse statement of Theorem 1, which yields an important characterization of the normal random vectors $\mathbf{X}_1$ and $\mathbf{X}_2$ when their dispersion matrices $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ are in Loewner order $\mathbf{\Sigma}_1 \leq \mathbf{\Sigma}_2$.

3 Stochastic characterization of Loewner dominance

We start by proving the following characterization theorem.

Theorem 2. *Let $\mathbf{X}_1 \sim N_k(\mathbf{0}, \mathbf{\Sigma}_1)$ and $\mathbf{X}_2 \sim N_k(\mathbf{0}, \mathbf{\Sigma}_2)$, $k \geq 1$, where $\mathbf{\Sigma}_1 > \mathbf{0}$. Then*

$$\mathrm{P}(\mathbf{X}_1 \in A) \geq \mathrm{P}(\mathbf{X}_2 \in A)$$

for all convex and symmetric (with respect to the origin) sets $A \subset \mathbf{R}^k$ iff $\mathbf{\Sigma}_1 \leq \mathbf{\Sigma}_2$.

Proof. In view of Theorem 1 we need only to prove the following assertion:

(i) If

$$\mathrm{P}(\mathbf{X}_1 \in A) \geq \mathrm{P}(\mathbf{X}_2 \in A) \tag{6}$$

holds for all convex and symmetric sets $A \subset \mathbf{R}^k$, then $\mathbf{\Sigma}_1 \leq \mathbf{\Sigma}_2$.

Since $\mathbf{\Sigma}_1$ is positive definite, then there exists a nonsingular matrix $\mathbf{Q}$ such that

$$\mathbf{Q}'\boldsymbol{\Sigma}_1\mathbf{Q} = \mathbf{I}_k \quad \text{and} \quad \mathbf{Q}'\boldsymbol{\Sigma}_2\mathbf{Q} = \mathbf{D}, \tag{7}$$

where $\mathbf{D} = \text{diag}(d_1, d_2, \ldots, d_k)$ and $d_1, d_2, \ldots, d_k$ are the eigenvalues of $\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_2$, all of which are positive (cf. Lancaster and Tismenetsky 1985, Theorem 2, p. 185). It follows from (7) that

$$\boldsymbol{\Sigma}_1 \leq \boldsymbol{\Sigma}_2 \quad \text{iff} \quad d_i \geq 1 \quad \text{for every } i = 1, 2, \ldots, k$$

and

$$\mathbf{Q}'\mathbf{X}_1 \sim N_k(\mathbf{0}, \mathbf{I}_k) \quad \text{and} \quad \mathbf{D}^{-1/2}\mathbf{Q}'\mathbf{X}_2 \sim N_k(\mathbf{0}, \mathbf{I}_k),$$

where $\mathbf{D}^{-1/2} = \text{diag}(d_1^{-1/2}, d_2^{-1/2}, \ldots, d_k^{-1/2})$. Let's denote for simplicity $\mathbf{Q}'\mathbf{X}_1 = \mathbf{Z}$ and $\mathbf{D}^{-1/2}\mathbf{Q}'\mathbf{X}_2 = \mathbf{V}$. Then it is clear that

$$\text{P}(\mathbf{X}_1 \in A) = \text{P}(\mathbf{Z} \in \mathbf{Q}'A)$$

and

$$\text{P}(\mathbf{X}_2 \in A) = \text{P}(\mathbf{Q}'\mathbf{X}_2 \in \mathbf{Q}'A) = \mathbf{P}(\mathbf{V} \in \mathbf{D}^{-1/2}\mathbf{Q}'A).$$

Since $\mathbf{Q}$ is nonsingular, the set $\mathbf{Q}'A$ is convex and symmetric iff A is convex and symmetric.

Suppose now for a moment that $\boldsymbol{\Sigma}_1 \leq \boldsymbol{\Sigma}_2$ *does not hold*, i.e. $d_i < 1$ for some value of i, say $d_1 < 1$. Let's choose now such a convex symmetric set $A \subset \mathbf{R}^k$ that

$$\mathbf{Q}'A = [-1, 1] \times \mathbf{R} \times \cdots \times \mathbf{R} \subset \mathbf{R}^k.$$

For such a set we have

$$\begin{aligned}
\text{P}(\mathbf{X}_1 \in A) &= \text{P}(\mathbf{Z} \in \mathbf{Q}'A) \\
&= \text{P}(Z_1 \in [-1, 1]) \\
&< \text{P}(V_1 \in d_1^{-1/2}[-1, 1]) \\
&= \text{P}(\mathbf{V} \in \mathbf{D}^{-1/2}\mathbf{Q}'A) \\
&= \text{P}(\mathbf{X}_2 \in A),
\end{aligned}$$

where Z_1 and V_1 are the first elements of $\mathbf{Z}$ and $\mathbf{V}$, respectively. Thus the assumption $d_1 < 1$ has led to the conclusion that the inequality (6) *does not hold* for all convex and symmetric sets $A \subset \mathbf{R}^k$. This proves the assertion (i). □

Remark 1. According to Eaton and Perlman (1991), $\mathbf{X}_1 \sim N_k(\mathbf{0}, \boldsymbol{\Sigma}_1)$ is said to be more concentrated than $\mathbf{X}_2 \sim N_k(\mathbf{0}, \boldsymbol{\Sigma}_2)$ iff $\boldsymbol{\Sigma}_1 \leq \boldsymbol{\Sigma}_2$. Hence Theorem 2 gives an equivalent characterization of '$\mathbf{X}_1$ more concentrated than $\mathbf{X}_2$'.

Motivated by Theorem 2, we define now a generalization of the DS-optimality criterion. We call it the SC-optimality criterion (S for 'stochastic' and C for 'convex').

Definition 2. Let $\hat{\boldsymbol{\beta}}_1 = \hat{\boldsymbol{\beta}}(\xi_1)$ and $\hat{\boldsymbol{\beta}}_2 = \hat{\boldsymbol{\beta}}(\xi_2)$ be the LSE's of $\boldsymbol{\beta}$ in (1) under the designs ξ_1 and ξ_2, respectively and let $\mathscr{A}$ be a class of convex symmetric (with respect to the origin) sets in $\mathbf{R}^k$.

(i) If

$$\mathrm{P}(\hat{\boldsymbol{\beta}}(\xi_1) - \boldsymbol{\beta} \in A) \geq \mathrm{P}(\hat{\boldsymbol{\beta}}(\xi_2) - \boldsymbol{\beta} \in A) \quad \text{for all } A \in \mathscr{A},$$

then the design ξ_1 dominates ξ_2 with respect to the $\mathrm{SC}_{\mathscr{A}}$-criterion.

(ii) If

$$\mathrm{P}(\hat{\boldsymbol{\beta}}(\xi_1) - \boldsymbol{\beta} \in A) \geq \mathrm{P}(\hat{\boldsymbol{\beta}}(\xi_2) - \boldsymbol{\beta} \in A)$$

for all convex symmetric (with respect to the origin) sets $A \subset \mathbf{R}^k$, then ξ_1 dominates ξ_2 with respect to the SC-criterion.

A design ξ^* for the LSE of $\boldsymbol{\beta}$ in (1) is $\mathrm{SC}_{\mathscr{A}}$*-optimal if*

$$\mathrm{P}(\hat{\boldsymbol{\beta}}(\xi^*) - \beta \in A) \geq \mathrm{P}(\hat{\boldsymbol{\beta}}(\xi) - \beta \in A) \quad \text{for all } A \in \mathscr{A}$$

and for all designs ξ. A design ξ^* is SC-optimal if it is $\mathrm{SC}_{\mathscr{A}}$-optimal for the class $\mathscr{A}$ of all convex symmetric (with respect to the origin) sets in $\mathbf{R}^k$.

Loewner dominance and SC-dominance induce by Theorem 2 the same partial ordering among designs. Thus, for example, the set of all admissible designs under Loewner dominance is equivalent to the set of all admissible designs under SC-dominance. Further, a design is Loewner optimal iff it is SC-optimal.

Theorem 2 also gives a sufficient condition for *Kiefer dominance* (see e.g. Pukelsheim 1993, Chapter 14). Indeed, let $\mathscr{H}$ be a subgroup of the set $\mathscr{C}$ of all orthogonal $k \times k$-matrices and suppose that the set of all available moment matrices is invariant with respect to $\mathscr{H}$. A design ξ_1 dominates ξ_2 in the Kiefer ($\mathscr{H}$) sense if there exists a matrix $\mathbf{A}$ belonging to the convex hull $\mathrm{conv}\{\mathbf{H}\mathbf{M}(\xi_2)\mathbf{H}', \mathbf{H} \in \mathscr{H}\}$ such that $\mathbf{M}(\xi_1) \geq \mathbf{A}$. It is easy to see that the smaller set $\mathscr{H}$ the stronger Kiefer ($\mathscr{H}$)-criterion. The smallest set $\mathscr{H}$ is the singleton set $\mathscr{H} = \{\pm \mathbf{I}_k\}$. But Kiefer ($\{\pm \mathbf{I}_k\}$)-criterion is simply Loewner criterion and it is stronger than Kiefer ($\mathscr{H}$)-criterion for any other $\mathscr{H}$.

By Lemma 1 the set of designs $\xi_p = \{a, b; p, 1-p\}$, $0 < p < 1$ is *a complete class* for the LSE of $\boldsymbol{\beta}$ in (2) relative to the Loewner dominance. It follows from Theorem 2 that the designs ξ_p, $0 < p < 1$, form a complete class also relative to SC-dominance. On the basis of equation (3) it is also clear that there is no Loewner optimal design, or equivalently, no SC-optimal design for the LSE of $\boldsymbol{\beta}$ in (2).

Further, the above results yield that any $\mathrm{SC}_{\mathscr{A}}$-optimal design is of the form $\xi_p = \{a, b; p, 1-p\}$. If we specialize on a certain class $\mathscr{A}$ of convex symmetric (with respect to the origin) sets in $\mathbf{R}^k$, there might be possible to find an optimal design. In fact, DS-optimality is a special case of $SC_{\mathscr{A}}$-optimality. If $\mathscr{A}$ is taken to be the class of all k-dimensional balls centered at the origin, then DS-optimality follows. We know that although there is no SC-optimal design for the LSE of $\boldsymbol{\beta}$ in (2) with $[a,b]=[-1,1]$, the design $\xi_{1/2}=\{-1,1;\frac{1}{2},\frac{1}{2}\}$ is the unique DS-optimal design.

The next result shows that we can extend the class of 2-dimensional balls centered at the origin so that the design $\xi_{1/2}=\{-1,1;\frac{1}{2},\frac{1}{2}\}$ still remains optimal for the LSE of $\boldsymbol{\beta}$ in (2) with $[a,b]=[-1,1]$.

Lemma 2. *The design $\xi_{1/2}=\{-1,1;\frac{1}{2},\frac{1}{2}\}$ is $SC_{\mathscr{A}}$-optimal for the LSE of $\boldsymbol{\beta}$ in (2) with $[a,b]=[-1,1]$, where $\mathscr{A}$ is a class of all convex symmetric (with respect to the axes) sets in $\mathbf{R}^2$.*

Proof. We utilize the considerations in the beginning of Section 2 and the proof of Lemma 1. Let's consider only the designs $\xi_p=\{-1,1;p,1-p\}$, $0<p<1$, having a moment matrix

$$\mathbf{M}(\xi_p)=\begin{pmatrix}1 & d\\ d & 1\end{pmatrix},\quad d=1-2p\in(-1,1).$$

It is easy to see that the eigenvalues of $\mathbf{M}(\xi_p)$ are $\lambda_{\min}=1-|d|$, $\lambda_{\max}=1+|d|$. Let $\mathbf{M}(\xi_p)=\mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}'$ be the spectral decomposition of $\mathbf{M}(\xi_p)$, where $\boldsymbol{\Lambda}=\mathrm{diag}[1-|d|,1+|d|]$ and $\mathbf{Q}$ is an orthogonal matrix. Define $\mathbf{Z}=\dfrac{\sqrt{n}}{\sigma}\boldsymbol{\Lambda}^{1/2}\mathbf{Q}'(\hat{\boldsymbol{\beta}}(\xi_p)-\boldsymbol{\beta})$ and note that $\mathbf{Z}\sim N_2(\mathbf{0},\mathbf{I}_2)$. Then

$$\begin{aligned}\mathrm{P}(\hat{\boldsymbol{\beta}}(\xi_p)-\boldsymbol{\beta}\in A)&=\mathrm{P}\left(\frac{\sigma}{\sqrt{n}}\mathbf{Q}\boldsymbol{\Lambda}^{-1/2}\mathbf{Z}\in A\right)\\&=\mathrm{P}\left(\left(\frac{Z_1}{\sqrt{1-|d|}},\frac{Z_2}{\sqrt{1+|d|}}\right)\in\frac{\sqrt{n}}{\sigma}\mathbf{Q}'A\right).\end{aligned}$$

It is easy to understand that the orthogonal matrix $\mathbf{Q}$ is either

$$\pm\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{pmatrix}\quad\text{or}\quad\pm\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{pmatrix}$$

if $d\neq0$. In any case the set $\dfrac{\sqrt{n}}{\sigma}\mathbf{Q}'A$ is convex, symmetric (with respect to the origin) and permutation-symmetric in $\mathbf{R}^2$. Hence, according to Theorem 7.4.6 of Tong (1990),

$$\mathrm{P}\left(\left(\frac{Z_1}{d_1},\frac{Z_2}{d_2}\right)\in\frac{\sqrt{n}}{\sigma}\mathbf{Q}'A\right)$$

is a Schur-concave function of $(\log d_1, \log d_2)$. Since $(\log 1, \log 1) \prec^w (\log\sqrt{1-|d|}, \log\sqrt{1+|d|})$, where $\prec^w$ means weak upper majorization (see Marshall and Olkin 1979, Definition A.2), we get

$$\mathrm{P}\left(\left(\frac{Z_1}{\sqrt{1-|d|}}, \frac{Z_2}{\sqrt{1+|d|}}\right) \in \frac{\sqrt{n}}{\sigma}\mathbf{Q}'A\right) \le \mathrm{P}\left(\mathbf{Z} \in \frac{\sqrt{n}}{\sigma}\mathbf{Q}'A\right)$$

$$= \mathrm{P}\left(\mathbf{Z} \in \frac{\sqrt{n}}{\sigma}A\right) = \mathrm{P}\big(\hat{\boldsymbol{\beta}}(\xi_{1/2}) - \boldsymbol{\beta} \in A\big). \quad \square$$

The next corollary is useful in searching for optimal designs. Given a design $\xi = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_l; p_1, p_2, \ldots, p_l\}$, *the reflected design* ξ^R is defined by $\xi^R = \{-\mathbf{x}_1, -\mathbf{x}_2, \ldots, -\mathbf{x}_l; p_1, p_2, \ldots, p_l\}$. The designs ξ and ξ^R have the same even moments, while the odd moments of ξ^R have a reversed sign. The moment matrix of ξ^R is given by $\mathbf{M}(\xi^R) = \mathbf{TM}(\xi)\mathbf{T}'$, where $\mathbf{T} = \mathrm{diag}(1, -1, 1, -1, \ldots, \pm 1)$ is a diagonal matrix with diagonal elements $1, -1, 1, -1, \ldots, \pm 1$.

Corollary 1. *Let $\mathscr{A}$ be a class of all convex symmetric (with respect to the axes) sets in $\mathbf{R}^2$. Then the following statements hold for any design ξ for the LSE of $\boldsymbol{\beta}$ in (2) with $[a, b] = [-1, 1]$:*

(i) $\mathrm{P}(\hat{\boldsymbol{\beta}}(\xi) - \boldsymbol{\beta} \in A) = \mathrm{P}(\hat{\boldsymbol{\beta}}(\xi^R) - \boldsymbol{\beta} \in A)$ *for all $A \in \mathscr{A}$;*
(ii) the symmetrized design $\bar{\xi} = \frac{1}{2}(\xi + \xi^R)$ dominates ξ with respect to the $\mathrm{SC}_{\mathscr{A}}$*-criterion.*

Proof. The first assertion is evident. Indeed, since $(\mathbf{TM}(\xi)\mathbf{T}')^{-1} = \mathbf{TM}(\xi)^{-1}\mathbf{T}'$, then

$$\mathrm{P}(\hat{\boldsymbol{\beta}}(\xi) - \boldsymbol{\beta} \in A) = \mathrm{P}(\mathbf{T}(\hat{\boldsymbol{\beta}}(\xi) - \boldsymbol{\beta}) \in \mathbf{T}A) = \mathrm{P}(\hat{\boldsymbol{\beta}}(\xi^R) - \boldsymbol{\beta} \in A)$$

due to the symmetricity of the set A with respect to the axes.

For the proof of the second assertion it is enough to note that since the moment matrix of ξ can be written as

$$\mathbf{M}(\xi) = \begin{pmatrix} 1 & h \\ h & c^2 \end{pmatrix}, \quad c > |h| \ge 0,$$

then

$$\mathrm{P}(\hat{\boldsymbol{\beta}}(\xi) - \boldsymbol{\beta} \in A) = \frac{\sqrt{c^2 - h^2}}{2\pi}\int_A e^{-\mathbf{x}'\mathbf{M}(\xi)\mathbf{x}/2}\, d\mathbf{x}$$

$$= \frac{\sqrt{1 - d^2}}{2\pi}\int_{A_c} e^{-\mathbf{x}'\tilde{\mathbf{M}}(\xi)\mathbf{x}/2}\, d\mathbf{x},$$

where

$$\tilde{\mathbf{M}}(\xi) = \begin{pmatrix} 1 & d \\ d & 1 \end{pmatrix}, \quad d = \frac{h}{c} \in (-1, 1),$$

$$A_c = \left\{(x, y) : \left(x, \frac{y}{c}\right) \in A\right\}.$$

From Lemma 2 it follows that

$$\begin{aligned} \mathrm{P}(\hat{\boldsymbol{\beta}}(\xi) - \boldsymbol{\beta} \in A) &= \frac{\sqrt{1 - d^2}}{2\pi} \int_{A_c} e^{-\mathbf{x}'\tilde{\mathbf{M}}(\xi)\mathbf{x}/2}\, d\mathbf{x} \le \frac{1}{2\pi} \int_{A_c} e^{-(\mathbf{x}'\mathbf{x})/2}\, d\mathbf{x} \\ &= \frac{c}{2\pi} \int_{A} e^{-\mathbf{x}'\mathbf{M}(\bar{\xi})\mathbf{x}/2}\, d\mathbf{x} \\ &= \mathrm{P}(\hat{\boldsymbol{\beta}}(\bar{\xi}) - \boldsymbol{\beta} \in A), \quad \text{where } \mathbf{M}(\bar{\xi}) = \begin{pmatrix} 1 & 0 \\ 0 & c^2 \end{pmatrix}. \quad \square \end{aligned}$$

The next example shows that the result of Lemma 2 is not necessarily true if the elements $A \in \mathbf{R}^2$ of $\mathscr{A}$ are convex and symmetric sets with respect to the origin but not with respect to the axes.

Example. Consider the set of rectangulars in $\mathbf{R}^2$ of the form

$$\mathscr{A} = \{\mathbf{C}A_c, c \ge 1\},$$

where $A_c = [-1, 1] \times [-c, c]$ and

$$\mathbf{C} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Evidently, each $A \in \mathscr{A}$ is convex and symmetric with respect to the origin but not with respect to the axes (see Figure 1).

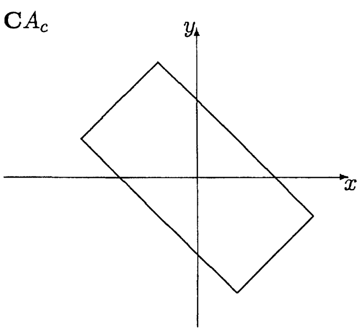

Fig. 1. A set from $\mathscr{A}$ is a rectangular that is convex and symmetric with respect to the origin but not with respect to the axes.

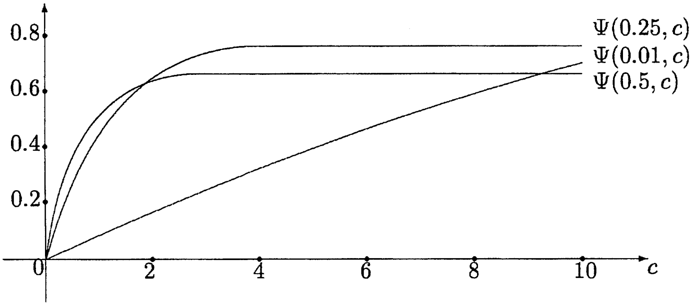

Fig. 2. The graph of $\Psi(p,c)$ for $p = 0.01, 0.25, 0.5$ and $\frac{n}{\sigma^2} = 1$.

We show that for the LSE of $\boldsymbol{\beta}$ in (2) with $[a,b] = [-1,1]$, the design $\xi_{1/2} = \{-1,1;\frac{1}{2},\frac{1}{2}\}$ is not SC$_{\mathscr{A}}$-optimal. Indeed, denote

$$\Psi(p,c) = \mathrm{P}(\hat{\boldsymbol{\beta}}(\xi_p) - \boldsymbol{\beta} \in \mathbf{C}A_c),$$

where $\xi_p = \{-1,1;p,1-p\}$. After simple calculations we get

$$\Psi(p,c) = \left[2\Phi\left(\frac{\sqrt{2n(1-p)}}{\sigma}\right) - 1\right]\left[2\Phi\left(\frac{c\sqrt{2np}}{\sigma}\right) - 1\right],$$

where Φ is the distribution function of the standard normal law.

Thus for any given $p \in (0,\frac{1}{2})$ there exists $c_0 = c_0(p) > 1$ such that

$$\Psi(\tfrac{1}{2},c) < \Psi(p,c) \qquad \forall c > c_0$$

and

$$\Psi(\tfrac{1}{2},c) > \Psi(p,c) \qquad \forall 0 < c < c_0.$$

The numerical values of c_0 for certain values of p in the case when $\frac{n}{\sigma^2} = 1$ are given in the following table:

p	0.01	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.49
c_0	9.32	4.25	3.08	2.55	2.20	1.95	1.73	1.54	1.36	1.19	1.04

In fact, the designs $\xi_p = \{-1,1;p,1-p\}$, $0 < p < \frac{1}{2}$ form a complete class, but there is no SC$_{\mathscr{A}}$-optimal design here.

4 Multifactor first degree polynomial models

Let us look at an m-way first-degree polynomial fit model

$$Y_{ij} = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_m x_{im} + E_{ij}, \quad i = 1,2,\ldots,l;\ j = 1,2,\ldots,n_i \tag{8}$$

with m regression variables and l experimental conditions $\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{im})'$, $i = 1, 2, \ldots, l$. We assume now that the experimental domain for the model (8) is an m-dimensional Euclidean ball of radius $\sqrt{m}$, that is $\mathcal{T}_{\sqrt{m}} = \{\mathbf{x} \in \mathbf{R}^m : \|\mathbf{x}\| \le \sqrt{m}\}$. Therefore, the regression range is of the form

$$\left\{ \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix} | \mathbf{x} \in \mathcal{T}_{\sqrt{m}} \right\} \subset \mathcal{T}_{\sqrt{m+1}}. \tag{9}$$

Denote $\mu_{0k} = \sum_{i=1}^{l} p_i x_{ik}$, $k = 1, 2, \ldots, m$ and $\mu_{jk} = \sum_{i=1}^{l} p_i x_{ij} x_{ik}$ for $j,\ k = 1, 2, \ldots, m$. Then the moment matrix of an l-point design

$$\xi = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_l; p_1, p_2, \ldots, p_l\} \tag{10}$$

is of the form

$$\mathbf{M}(\xi) = \sum_{i=1}^{l} p_i \begin{pmatrix} 1 \\ \mathbf{x}_i \end{pmatrix} (1, \mathbf{x}_i) = \begin{pmatrix} 1 & \mu_{01} & \cdots & \mu_{0m} \\ \mu_{10} & \mu_{11} & \cdots & \mu_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{m0} & \mu_{m1} & \cdots & \mu_{mm} \end{pmatrix}.$$

Consider an l-point design (10) which is in the range (9), $l \ge m + 1$. Let

$$\lambda_1 \mathbf{w}_1 \mathbf{w}_1' + \lambda_2 \mathbf{w}_2 \mathbf{w}_2' + \cdots + \lambda_{m+1} \mathbf{w}_{m+1} \mathbf{w}_{m+1}' = \mathbf{M}(\xi) \tag{11}$$

be the spectral decomposition of $\mathbf{M}(\xi)$, where $\mathbf{w}_i$ and $\lambda_i > 0$ are orthonormal eigenvectors and the eigenvalues of $\mathbf{M}(\xi)$, respectively. Note that

$$\mathrm{tr}[\mathbf{M}(\xi)] = \lambda_1 + \lambda_2 + \cdots + \lambda_{m+1} = \sum_{i=1}^{l} p_i (1 + \mathbf{x}_i' \mathbf{x}_i) \le m + 1, \tag{12}$$

since by assumption $\mathbf{x}_i \in \mathcal{T}_{\sqrt{m}}$ for all $i = 1, 2, \ldots, l$.

Denote $\tilde{\mathbf{x}}_i = \sqrt{\lambda_1 + \lambda_2 + \cdots + \lambda_{m+1}}\,\mathbf{w}_i$ and $r_i = \dfrac{\lambda_i}{\lambda_1 + \lambda_2 + \cdots + \lambda_{m+1}}$, $i = 1, 2, \ldots, m + 1$, and consider the $(m + 1)$-point design

$$\tilde{\xi} = \{\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \ldots, \tilde{\mathbf{x}}_{m+1}; r_1, r_2, \ldots, r_{m+1}\}.$$

Clearly, $\tilde{\mathbf{x}}_i \in \mathcal{T}_{\sqrt{m+1}}$, $i = 1, 2, \ldots, m + 1$, and the designs $\tilde{\xi}$ and ξ have the same information matrix, i.e. $\tilde{\xi}$ and ξ are information equivalent designs, though $\tilde{\xi}$ is not in the range (9). Thus for any l-point design ξ from the regression range (9) for the LSE of $\boldsymbol{\beta}$ in (8) there exists an information equivalent $(m + 1)$-point design $\tilde{\xi}$ on $\mathcal{T}_{\sqrt{m+1}}$ with orthogonal support vectors. We say that $\tilde{\xi}$ is an *orthogonal design*.

Before proceeding to Theorem 3 we introduce two further design notions that will be needed. A design $\xi_0 = \{\mathbf{x}_1^0, \mathbf{x}_2^0, \ldots, \mathbf{x}_l^0; p_1, p_2, \ldots, p_l\}$ is a *rotation* of the design $\xi = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_l; p_1, p_2, \ldots, p_l\}$ if there exists an orthogonal matrix $\mathbf{C}$ such that $\mathbf{x}_i^0 = \mathbf{C}\mathbf{x}_i$, $i = 1, 2, \ldots, l$. The moment matrix $\mathbf{M}(\xi_0)$ of ξ_0 is $\mathbf{C}\mathbf{M}(\xi)\mathbf{C}'$. Let $\mathbf{x}_i \in \mathscr{T}_{\sqrt{m}}$, $i = 1, 2, \ldots, m+1$, fulfill the conditions

$$1 + \mathbf{x}_i'\mathbf{x}_i = 1 + m, \quad 1 + \mathbf{x}_i'\mathbf{x}_j = 0 \tag{13}$$

for all $i \neq j \leq m+1$. Then the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{m+1}$ satisfying (13) span a convex body in $\mathbf{R}^m$ called a *regular simplex*. A design $\xi = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{m+1}; p_1, p_2, \ldots, p_{m+1}\}$ which places weights p_i, $i = 1, 2, \ldots, m+1$, on the vertices of a regular simplex in $\mathbf{R}^m$ is a *simplex design* (cf. Pukelsheim 1993, p. 391).

Now we prove that any design ξ of the form (10) from the regression range (9) can be dominated in the Loewner sense by a rotation of a simplex design.

Theorem 3. *Let ξ be an l-point design on the regression range (9) for the LSE of $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_m)'$ in (8), $l \geq m+1$. Then there exists such a rotation $\hat{\xi}$ of an $(m+1)$-point simplex design ξ_0 that $\hat{\xi}$ dominates ξ in the Loewner sense, i.e. $\hat{\xi} \succ_L \xi$. Equivalently, $\boldsymbol{\lambda}(\mathbf{M}(\xi_0)) \geq \boldsymbol{\lambda}(\mathbf{M}(\xi))$, where $\boldsymbol{\lambda}(\mathbf{M})$ is the vector of ordered eigenvalues of $\mathbf{M}$ and $\geq$ refers to the usual entrywise ordering.*

Proof. Let any l-point design ξ with $l \geq m+1$ on the regression range (9) be given. Define an $(m+1)$-point orthogonal design $\hat{\xi} = \{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots, \hat{\mathbf{x}}_{m+1}; r_1, r_2, \ldots, r_{m+1}\}$ on the boundary of $\mathscr{T}_{\sqrt{m+1}}$ as follows:

$$\hat{\mathbf{x}}_i = \sqrt{m+1}\,\mathbf{w}_i, \quad r_i = \frac{\lambda_i}{\lambda_1 + \lambda_2 + \cdots + \lambda_{m+1}},$$

where $\mathbf{w}_i$ and $\lambda_i > 0$ are from (11). It is easy to see that $\hat{\xi}$ dominates ξ in the Loewner sense since

$$\mathbf{M}(\hat{\xi}) = \sum_{i=1}^{m+1} r_i \hat{\mathbf{x}}_i \hat{\mathbf{x}}_i' = \frac{m+1}{\lambda_1 + \lambda_2 + \cdots + \lambda_{m+1}} \mathbf{M}(\xi) \geq \mathbf{M}(\xi)$$

by (12) and consequently $\hat{\xi} \succ_L \xi$, though again $\hat{\xi}$ is not in the range (9).

Finally, it remains to show that there exists an $(m+1)$-point simplex design ξ_0 which is a rotation of $\hat{\xi}$. Indeed, denote $\mathbf{Q} = (\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{m+1})$ and let $\mathbf{C}_0$ be an orthogonal matrix whose first row is $(1/\sqrt{m+1}, \ldots, 1/\sqrt{m+1})$. Then the design $\xi_0 = \{\mathbf{x}_1^0, \mathbf{x}_2^0, \ldots, \mathbf{x}_{m+1}^0; r_1, r_2, \ldots, r_{m+1}\}$ with

$$(\mathbf{x}_1^0, \mathbf{x}_2^0, \ldots, \mathbf{x}_{m+1}^0) = \mathbf{C}_0\mathbf{Q}'(\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots, \hat{\mathbf{x}}_{m+1}) = \sqrt{m+1}\,\mathbf{C}_0$$

is a rotation of $\hat{\xi}$ (and $\hat{\xi}$ is a rotation of ξ_0), and the support points $\mathbf{x}_1^0, \mathbf{x}_2^0, \ldots, \mathbf{x}_{m+1}^0$ of ξ_0 belong to the range (9) and fulfill the conditions (13), i.e. ξ_0 is a simplex design. This concludes the proof of the theorem. □

The above results have important implications since all comparisons of designs are based on moment matrices and a reasonable *optimality criterion* is *isotonic* with respect to the Loewner ordering. An optimality criterion is a function ϕ from the closed cone of nonnegative definite matrices into the real line (cf. Pukelsheim 1993, p. 114). Relative to the criterion ϕ, a design ξ_1 is *at* *least as good as* another design ξ_2 when $\phi(\mathbf{M}_1) \geq \phi(\mathbf{M}_2)$, where $\mathbf{M}_1$ and $\mathbf{M}_2$ are the moment matrices of ξ_1 and ξ_2, respectively. A criterion ϕ is isotonic with respect to Loewner ordering if

$$\mathbf{M}_1 \geq \mathbf{M}_2 \geq \mathbf{0} \Rightarrow \phi(\mathbf{M}_1) \geq \phi(\mathbf{M}_2).$$

The matrix means are the most prominent optimality criteria and they enjoy many desired properties like isotonicity (cf. Shah and Sinha 1989, Chapter 1, or Pukelsheim 1993, p. 119 and Chapter 6). The classical A-, E-, D- and T-optimality criteria are just particular cases of matrix means. Also DS-optimality criterion is isotonic (see (5)).

Theorem 3 yields immediately the following corollary.

Corollary 2. *For any l-point design ξ, $l \geq m+1$, for the LSE of $\boldsymbol{\beta}$ in (8) there exists an $(m+1)$-point simplex design that dominates ξ with respect to any optimality criterion which is isotonic with respect to the Loewner ordering.*

Corollary 2 says that e.g. for the SC-optimality criterion the class of all simplex designs for the LSE of $\boldsymbol{\beta}$ in (8) is complete though it does not guarantee the existence of a SC-optimal design. A design which places uniform weight $1/(m+1)$ on the vertices $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{m+1}$ of a regular simplex in $\mathbf{R}^m$ is called a *uniform simplex design*. Any rotation of a uniform simplex design is a uniform simplex design. Liski et al. (1999) showed that an $(m+1)$-point design ξ for the LSE of $\boldsymbol{\beta}$ in (8) is DS-optimal iff it is a uniform simplex design.

Theorem 3 also proves Kiefer dominance of simplex designs for the LSE of $\boldsymbol{\beta}$ in (8). Indeed, it follows from Theorem 3 that any l-point design ξ, $l \geq m+1$, is dominated in the Kiefer ($\mathscr{C}$) sense by an $(m+1)$-point simplex design ξ_0.

Consider also an m-way first degree model without a constant term,

$$Y_{ij} = \beta_1 x_{i1} + \cdots + \beta_m x_{im} + E_{ij}, \quad i = 1, 2, \ldots, l;\ j = 1, 2, \ldots, n_i. \tag{14}$$

Again assume that the experimental domain is $\mathscr{T}_{\sqrt{m}}$. Repeating the considerations leading to Theorem 3 and Corollary 2 we get the corresponding results for the model (14).

Theorem 4. *Let ξ be an l-point design on $\mathscr{T}_{\sqrt{m}}$ for the LSE of $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_m)'$ in (14), $l \geq m$. Then there exists an m-point orthogonal design $\hat{\xi}$ that dominates ξ in the Loewner sense, i.e. $\hat{\xi} \succ_L \xi$.*

Corollary 3. *For any l-point design ξ, $l \geq m$, for the LSE of $\boldsymbol{\beta}$ in (14) there exist an m-point orthogonal design that dominates ξ with respect to the SC-optimality criterion.*

It is again clear that there is no Loewner optimal or SC-optimal m-point orthogonal design for the LSE of $\boldsymbol{\beta}$ in (8). Liski et al. (1999) showed that a DS-optimal design on $\mathscr{T}_{\sqrt{m}}$ for the LSE of $\boldsymbol{\beta}$ in (14) always exists. Thus taking as a class of convex sets $\mathscr{A}$ a class of all m-dimensional balls centered at the origin yields an optimal design with respect to the $\mathrm{SC}_{\mathscr{A}}$-optimality criterion.

5 Generalizations

Up to now we have assumed that the observations follow a multivariate normal distribution. However, the corresponding results hold also for an elliptically contoured distribution. A random vector $\mathbf{X}$ has an elliptically contoured distribution, written $\mathbf{X} \sim ECD(g, \mathbf{a}, \mathbf{\Sigma})$, if its density function is of the form

$$f(\mathbf{x}) = |\mathbf{\Sigma}|^{-1/2} g[(\mathbf{x} - \mathbf{a})' \mathbf{\Sigma}^{-1} (\mathbf{x} - \mathbf{a})],$$

where $g : R \to [0, \infty)$ is nonincreasing and $D(\mathbf{X}) = \mathbf{\Sigma} > \mathbf{0}$.

Theorem 1 and Theorem 2 can be formulated for elliptically contoured distributions as well.

We generalize now Theorem 2 for random vectors $\mathbf{X}_1 \sim ECD(g, \mathbf{0}, \mathbf{\Sigma}_1)$ and $\mathbf{X}_2 \sim ECD(g, \mathbf{0}, \mathbf{\Sigma}_2)$ with $\mathbf{\Sigma}_2 \geq \mathbf{\Sigma}_1 > \mathbf{0}$. The corresponding generalization of Theorem 1 is obvious.

Theorem 5. *Let* $\mathbf{X}_1 \sim ECD(g, \mathbf{0}, \mathbf{\Sigma}_1)$ *and* $\mathbf{X}_2 \sim ECD(g, \mathbf{0}, \mathbf{\Sigma}_2)$ *be* $k \times 1$ *random vectors,* $k \geq 1$*, where* $\mathbf{\Sigma}_1 > \mathbf{0}$*. Then*

$$\mathrm{P}(\mathbf{X}_1 \in A) \geq \mathrm{P}(\mathbf{X}_2 \in A)$$

holds for all convex and symmetric (with respect to the origin) sets $A \subset \mathbf{R}^k$ *iff* $\mathbf{\Sigma}_1 \leq \mathbf{\Sigma}_2$.

The proof of Theorem 5 is similar to the proof of Theorem 2, except that a normal distribution is replaced by an elliptically contoured distribution.

Remark 2. Theorem 5 gives a useful characterization of peakedness in the sense of definition by Sherman (1955). Indeed, for $k \times 1$ random vectors $\mathbf{X}_1$ and $\mathbf{X}_2$, $\mathbf{X}_1$ is said to be more peaked than $\mathbf{X}_2$ if

$$\mathrm{P}(\mathbf{X}_1 \in A) \geq \mathrm{P}(\mathbf{X}_2 \in A)$$

for all convex and symmetric (with respect to the origin) sets $A \subset \mathbf{R}^k$. So, for $\mathbf{X}_1 \sim ECD(g, \mathbf{0}, \mathbf{\Sigma}_1)$ and $\mathbf{X}_2 \sim ECD(g, \mathbf{0}, \mathbf{\Sigma}_2)$, $\mathbf{X}_1$ is more peaked than $\mathbf{X}_2$ iff $\mathbf{\Sigma}_2 \geq \mathbf{\Sigma}_1 > \mathbf{0}$. As we see, for multivariate normal vectors 'more peaked' and 'more concentrated' are equivalent notions.

Remark 3. Following the proof of Theorem 2 it is easy to see that the converse statement of Theorem 1 holds even in more general setting.

Corollary 4. *Let* $\mathbf{X} = (X_1, \ldots, X_k)'$ *be such a random vector that* $\mathrm{E}(\mathbf{X}) = \mathbf{0}$, $\mathrm{D}(\mathbf{X}) = \mathbf{\Sigma} > \mathbf{0}$*, and the support of its distribution contains the origin as an interior point. If there exists a nonsingular* $k \times k$*-matrix* $\mathbf{F}$ *such that the inequality*

$$\mathrm{P}(\mathbf{X} \in A) \geq \mathrm{P}(\mathbf{F}\mathbf{X} \in A)$$

holds for all convex and symmetric sets $A \subset \mathbf{R}^k$*, then* $\mathbf{F}\mathbf{\Sigma}\mathbf{F}' \geq \mathbf{\Sigma}$.

Acknowledgements. The second author is very grateful to Prof. M. D. Perlman for fruitful discussions concerning Theorem 2. Both authors would also like to express their gratitude to the referees for careful reading of the manuscript and valuable remarks which improved an earlier version of the paper. Our thanks go also to Sven-Oliver Troschke for providing us with a copy of Stępniak and Otachel (1994).

References

Anderson TW (1955) The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. Proceedings of the American Mathematical Society 6:170–176

De la Garza A (1954) Spacing of information in polynomial estimation. Annals of Mathematical Statistics 25:123–130

Eaton ML and Perlman MD (1991) Concentration inequalities for multivariate distributions: I. Multivariate normal distributions. Statist. Probab. Lett. 12:487–504

Giovagnoli A and Wynn HP (1995) Multivariate dispersion orderings. Statist. Probab. Lett. 22:325–332

Hwang JT (1985) Universal domination and stochastic domination: estimation simultaneously under a broad class of loss functions. Annals of Statistics 13:295–314

Lancaster P and Tismenetsky M (1985) The theory of matrices. Academic Press, New York

Liski EP, Luoma A, Mandal NK and Sinha, Bikas K (1998) Pitman nearness, distance criterion and optimal regression designs. Calcutta Statistical Association Bulletin 48(191–192):179–194

Liski EP, Luoma A and Zaigraev A (1999) Distance optimality design criterion in linear models. Metrika 49:193–211

Marshall AW and Olkin I (1979) Inequalities: Theory of majorization and its applications. Academic Press, New York

Perlman MD (1989) TW Anderson's theorem on the integral of a symmetric unimodal function over a symmetric convex set and its applications in probability and statistics. In: The collected papers of TW Anderson: 1943–1985 (edited by George P. H. Styan). Vol. 2, 1627–1641

Pukelsheim F (1993) Optimal design of experiments. Wiley, New York

Shah KR and Sinha, Bikas K (1989) Theory of optimal designs. Springer-Verlag Lecture Notes in Statistics Series, No. 54

Sherman S (1955) A theorem on convex sets with applications. Annals of Mathematical Statistics 26:763–766

Sinha, Bikas K (1970) On the optimality of some design. Calcutta Statistical Association Bulletin 20:1–20

Stępniak C (1989) Stochastic ordering and Schur-convex functions in comparison of linear experiments. Metrika 36:291–298

Stępniak C and Otachel Z (1994) Some directions in comparison of linear experiments: a review. In: Proceedings of the International Conference LINSTAT'93, 239–244

Tong YL (1990) The multivariate normal distribution. Springer-Verlag, New York

Torgersen E (1991) Comparison of statistical experiments. Cambridge University Press, Cambridge