Metrika (2002) 56: 259–273

Metrika

Shape optimal design criterion in linear models

Alexander Zaigraev

Faculty of Mathematics and Computer Science, Nicholas Copernicus University, Chopin str. 12/18, PL-87-100 Toruń, Poland (e-mail: alzaig@mat.uni.torun.pl)

Abstract. Within the framework of classical linear regression model optimal design criteria of stochastic nature are considered. The particular attention is paid to the shape criterion. Also its limit behaviour is established which generalizes that of the distance stochastic optimality criterion. Examples of the limit maximin criterion are considered and optimal designs for the line fit model are found.

Key words: classical linear regression model, line fit model, shape and distance stochastic optimality criteria, limit maximin criterion

1 Introduction

There exists an extensive literature on optimal design criteria. For references see Shah and Sinha (1989) and Pukelsheim (1993), for example. The most exhaustive investigations have been carried out in the case of so-called traditional criteria like A-, D- or E-optimality. Also more sophisticated criteria like universal or Kiefer optimality are occupied a certain place in the literature. However, *stochastic optimality criteria* have not drawn much attention hitherto. This paper is an attempt to fill the gap, in some sense. In the next section we explain what we mean saying 'stochastic optimality criteria'.

In the paper we consider the classical linear regression 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 $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)'$ follows a multivariate normal distribution, $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)'$ is the $n \times k$ model matrix of the full rank k, $k \leq n$, $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_k)'$ 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 $\sigma^2 = \mathrm{V}(Y_i) > 0$ is unknown while $\mathbf{I}_n$ is the $n \times n$ identity matrix.

Let $\hat{\boldsymbol{\beta}}$ be the least squares estimator (LSE) of $\boldsymbol{\beta}$ being at the same time the best linear unbiased estimator. The dispersion matrix of $\hat{\boldsymbol{\beta}}$ is $\mathrm{D}(\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$.

In the sequel we deal with so-called *continuous designs*. Each continuous design ξ is a discrete probability measure taking values $p_i \geq 0$ at vectors $\mathbf{x}_i$, $i = 1, 2, \dots, l$, that is

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

The *moment matrix* of a design ξ is defined by $\mathbf{M}(\xi) = \sum_{i=1}^{l} p_i \mathbf{x}_i \mathbf{x}_i'$. If $p_i = \frac{n_i}{n}$, $i = 1, 2, \dots, l$, $l \leq n$, where n_i are integers and $\sum_{i=1}^{l} n_i = n$, then $\mathrm{D}(\hat{\boldsymbol{\beta}}) = \frac{\sigma^2}{n}\mathbf{M}^{-1}$. Throughout the paper, we write $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}(\xi)$ or $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}(\mathbf{M})$ to emphasize the dependence of $\hat{\boldsymbol{\beta}}$ from the design ξ or from the moment matrix $\mathbf{M}$, respectively.

In the paper we also refer to a line fit model when we have $n \geq 2$ uncorrelated responses

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

with expectations and variances $\mathrm{E}(Y_{ij}) = \beta_1 + \beta_2 x_i$ and $\mathrm{V}(Y_{ij}) = \sigma^2$, respectively. In this case a continuous design ξ specifies distinct values $x_1, x_2, \dots, x_l$ chosen from a given experimental domain (usually an interval $[a, b]$) and assigns to them weights $p_i \geq 0$ such that $\sum_{i=1}^{l} p_i = 1$. Here

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

The paper is organized as follows. Stochastic optimality criteria and the shape criterion in particular are discussed in Section 2. In Section 3 we collect examples of the limit maximin criterion and optimal designs for the line fit model (2). In Section 4 concluding remarks are given while the proof of Theorem 2 can be found in Appendix.

2 Shape optimal design criterion

An optimality criterion is a function from the closed cone of nonnegative definite matrices into the real line. Saying 'stochastic optimality criteria' we mean functions depending on the moment matrices through a probability. A typical example is the criterion

$$\mathrm{P}(\hat{\boldsymbol{\beta}}(\mathbf{M}) - \boldsymbol{\beta} \in A) \to \max \quad \forall A \in \mathscr{A}, \tag{3}$$

where $\mathscr{A}$ is a given class of subsets from $\mathbf{R}^k$. In the sequel we deal with criteria of type (3).

Of course, the terminology is rather relative. It is due to Sinha (1970) who introduced the concept of the *distance stochastic* (DS) *criterion* in certain treatment design settings. Liski et al. (1998, 1999) studied the properties of this criterion under the classical linear regression model (1).

Definition 1. *A design ξ^* is said to be DS(ε)-optimal for the LSE of $\boldsymbol{\beta}$ in (1) if it maximizes the probability $\mathrm{P}(\|\hat{\boldsymbol{\beta}}(\xi) - \boldsymbol{\beta}\| \le \varepsilon)$ for a given $\varepsilon > 0$, where $\|\cdot\|$ denotes the Euclidean norm in $\mathbf{R}^k$. When ξ^* is DS(ε)-optimal for all $\varepsilon > 0$, we say that ξ^* is DS-optimal.*

Clearly, the DS-criterion is that of type (3) with $\mathscr{A}$ to be a class of all k-dimensional balls centered at the origin.

The DS(ε)-criterion function is defined as

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

It is worth noting that the DS(ε)-optimal design itself is not of great interest from the viewpoint of practice since usually it depends on unknown σ.

The criterion is isotonic relative to Loewner ordering (cf. Liski et al. 1999), that is for any two moment matrices $\mathbf{M}_1$ and $\mathbf{M}_2$,

$$\mathbf{M}_1 \ge \mathbf{M}_2 \Rightarrow \psi_\varepsilon[\mathbf{M}_1] \ge \psi_\varepsilon[\mathbf{M}_2] \quad \forall \varepsilon > 0.$$

Here the relation $\mathbf{M}_1 \ge \mathbf{M}_2$ for two matrices means that $\mathbf{M}_1 - \mathbf{M}_2$ is a non-negative definite matrix.

Liski et al. (1999, Theorem 5.1) also studied the behaviour of the DS(ε)-criterion, when ε approaches 0 and ∞. These limiting cases have an interesting relationship with the traditional D- and E-optimality criteria. It turns out that the DS(ε)-criterion is *equivalent* to the D-criterion as $\varepsilon \to 0$ and to the E-criterion as $\varepsilon \to \infty$, that is:

(a) if $\psi_\varepsilon[\mathbf{M}_1] \ge \psi_\varepsilon[\mathbf{M}_2]$ for all sufficiently small $\varepsilon > 0$, then $\det \mathbf{M}_1 \ge \det \mathbf{M}_2$; if $\det \mathbf{M}_1 > \det \mathbf{M}_2$, then $\psi_\varepsilon[\mathbf{M}_1] > \psi_\varepsilon[\mathbf{M}_2]$ for all sufficiently small $\varepsilon > 0$;
(b) if $\psi_\varepsilon[\mathbf{M}_1] \ge \psi_\varepsilon[\mathbf{M}_2]$ for all sufficiently large ε, then $\lambda_{\min}(\mathbf{M}_1) \ge \lambda_{\min}(\mathbf{M}_2)$; if $\lambda_{\min}(\mathbf{M}_1) > \lambda_{\min}(\mathbf{M}_2)$, then $\psi_\varepsilon[\mathbf{M}_1] > \psi_\varepsilon[\mathbf{M}_2]$ for all sufficiently large ε.

In Section 3 we generalize this result. Another motivation for considering stochastic optimality criteria can be found in Liski and Zaigraev (2001). It turns out that the distance stochastic criterion is closely connected with such notions as *stochastic ordering* (cf. Marshall and Olkin 1979, Section 17) or *stochastic precision* (cf. Stępniak 1989), *universal domination* (cf. Hwang 1985), *D-ordering* (cf. Giovagnoli and Wynn 1995). Moreover, criteria of type (3) have an obvious relation with *peakedness* (cf. Sherman 1955), *concentration* (cf. Eaton and Perlman 1991) and, as it was shown in Liski and Zaigraev (2001), with *Loewner optimality* (cf. Pukelsheim 1993, Section 4).

Liski and Zaigraev (2001) also suggested a natural generalization of the DS-criterion. It was called the *stochastic convex* (SC) *criterion*.

Definition 2. *Let $\mathscr{A}$ be a class of subsets from $\mathbf{R}^k$ which are convex and symmetric with respect to the origin. A design ξ^* is said to be* $\mathrm{SC}_{\mathscr{A}}$*-optimal for the LSE of $\boldsymbol{\beta}$ in (1) if it maximizes the probability* $\mathrm{P}(\hat{\boldsymbol{\beta}}(\xi) - \boldsymbol{\beta} \in A)$ *for all $A \in \mathscr{A}$.*

A design ξ^* is SC*-optimal if it maximizes the probability* $\mathrm{P}(\hat{\boldsymbol{\beta}}(\xi) - \boldsymbol{\beta} \in A)$ *for all convex and symmetric sets $A \subset \mathbf{R}^k$ (symmetric with respect to the origin).*

Evidently, the $\mathrm{SC}_{\mathscr{A}}$-criterion with $\mathscr{A}$ to be a class of all k-dimensional balls centered at the origin is the DS-criterion. Another example of the $\mathrm{SC}_{\mathscr{A}}$-criterion is that when $\mathscr{A}$ is a class of all subsets from $\mathbf{R}^k$ which are convex and symmetric with respect to the axes. Due to Lemma 2 from Liski and Zaigraev (2001), the $\mathrm{SC}_{\mathscr{A}}$-optimal design for the LSE of $\boldsymbol{\beta}$ in (2) on the interval $[-1, 1]$ is of the form $\{-1, 1; 0.5, 0.5\}$, although there is no SC-optimal design here. Liski and Zaigraev (2001) also proved isotonicity of the SC-criterion relative to Loewner ordering.

From now on, we deal with a case when the class $\mathscr{A}$ in (3) is of the form

$$\mathscr{A} = \{\varepsilon A, \varepsilon > 0\},$$

where $A \subset \mathbf{R}^k$ is a given bounded set having the origin as an interior point. In particular, if

$$A = A_\rho = \{\mathbf{x} \in \mathbf{R}^k : \mathbf{x} = t\mathbf{e}, 0 \le t \le \rho(\mathbf{e}), \mathbf{e} \in \mathbf{S}^{k-1}\}, \tag{4}$$

where ρ is a positive continuous function on the unit sphere $\mathbf{S}^{k-1}$ in $\mathbf{R}^k$, then we call such a criterion the *shape stochastic* (SS) *criterion*. The function ρ is called the shape function since it determines the shape of the set A_ρ.

Definition 3. *A design ξ^* is said to be* $\mathrm{SS}_\rho(\varepsilon)$*-optimal for the LSE of $\boldsymbol{\beta}$ in (1) if it maximizes the probability* $\mathrm{P}(\hat{\boldsymbol{\beta}}(\xi) - \boldsymbol{\beta} \in \varepsilon A_\rho)$ *for a given $\varepsilon > 0$ and given ρ. When ξ^* is* $\mathrm{SS}_\rho(\varepsilon)$*-optimal for all $\varepsilon > 0$, we say that ξ^* is* SS_ρ*-optimal.*

The $\mathrm{SS}_\rho(\varepsilon)$-criterion function is defined as

$$\varphi_{\rho,\varepsilon}[\mathbf{M}] = \mathrm{P}(\hat{\boldsymbol{\beta}}(\mathbf{M}) - \boldsymbol{\beta} \in \varepsilon A_\rho).$$

Since

$$\hat{\boldsymbol{\beta}}(\mathbf{M}) - \boldsymbol{\beta} \sim \mathrm{N}_k\left(\mathbf{0}, \frac{\sigma^2}{n}\mathbf{M}^{-1}\right),$$

that function also can be rewritten in the form

$$\varphi_{\rho,\varepsilon}[\mathbf{M}] = \mathrm{P}\left(\frac{\sigma}{\sqrt{n}}\mathbf{M}^{-1/2}\mathbf{Z} \in \varepsilon A_\rho\right) = \mathrm{P}(\mathbf{Z} \in \delta \mathbf{M}^{1/2} A_\rho), \tag{5}$$

where $\delta = \dfrac{\sqrt{n}\varepsilon}{\sigma}$ and $\mathbf{Z} \sim \mathrm{N}_k(\mathbf{0}, \mathbf{I}_k)$. Again it should be noted that the $\mathrm{SS}_\rho(\varepsilon)$-optimal design, in general, depends on unknown σ.

As it follows from Liski and Zaigraev (2001), the SS_ρ-criterion is isotonic relative to Loewner ordering if and only if A_ρ is convex and symmetric with respect to the origin.

Next, consider several examples of the shape functions.

Examples.

1. If $\rho(\mathbf{e}) \equiv 1$ then A_ρ is a ball centered at the origin and the SS_ρ-criterion is simply the DS-criterion.

2. Let $\rho(\mathbf{e}) = (\mathbf{e}'\mathbf{D}\mathbf{e})^{-1/2}$, where $\mathbf{D}$ is a symmetric positive definite matrix. Then A_ρ is an ellipsoid centered at the origin. If $\mathbf{D}$ is diagonal then A_ρ is symmetric with respect to the axes otherwise it is symmetric only with respect to the origin.

3. If $\rho(\mathbf{e}) = \min\{a_1|e_1|^{-1}, a_2|e_2|^{-1}, \dots, a_k|e_k|^{-1}\}$, $a_i > 0$, $i = 1, 2, \dots, k$, then A_ρ is a parallelepiped centered at the origin and symmetric with respect to the axes.

Relevant $\mathrm{SS}_\rho(\varepsilon)$-criteria are considered in the next section.

The limit behaviour of the $\mathrm{SS}_\rho(\varepsilon)$-criterion is described in the next two theorems.

Theorem 1. *The* $\mathrm{SS}_\rho(\varepsilon)$*-criterion is equivalent to the D-criterion as* $\varepsilon \to 0$.

The proof is evident since

$$\varphi_{\rho,\varepsilon}[\mathbf{M}] = \frac{\delta^k (\det \mathbf{M})^{1/2}}{(2\pi)^{k/2}} \int_{A_\rho} e^{-(\delta^2/2)\mathbf{z}'\mathbf{M}\mathbf{z}}\, d\mathbf{z}$$

and the function under the integral approaches 1 as $\varepsilon \to 0$ while the integral over A_ρ approaches $v(A_\rho) > 0$, where $v(A_\rho)$ denotes the Lebesque measure (volume) of the set A_ρ.

It is easy to understand that the assertion of Theorem 1 also holds in more general case, namely for any set A with $0 < v(A) < \infty$.

An immediate useful consequence of Theorem 1 is the following. If the SS_ρ-optimal design exists then it is necessary also D-optimal.

Now let $\varepsilon \to \infty$. The particular case of the $\mathrm{SS}_\rho(\varepsilon)$-criterion when $\rho(\mathbf{e}) \equiv 1$ was considered in Liski et al. (1999). Here we give more general result.

Let ∂A_ρ be the boundary of the set A_ρ.

Theorem 2. *Let* $\rho(\mathbf{e})$ *be a twice continuously differentiable function in a neighbourhood of any point from the set*

$$\operatorname{Arg} \min_{\mathbf{e} \in \mathbf{S}^{k-1}} \rho^2(\mathbf{e})\mathbf{e}'\mathbf{M}\mathbf{e}$$

for any given $\mathbf{M}$*. Then the* $\mathrm{SS}_\rho(\varepsilon)$*-criterion is equivalent to the maximin criterion*

$$\min_{\mathbf{x} \in \partial A_\rho} \mathbf{x}'\mathbf{M}\mathbf{x} = \min_{\mathbf{e} \in \mathbf{S}^{k-1}} \rho^2(\mathbf{e})\mathbf{e}'\mathbf{M}\mathbf{e} \to \max$$

as $\varepsilon \to \infty$.

The condition of Theorem 2 means that ∂A_ρ should be sufficiently smooth. But it does not require the boundary to be smooth at any point. Indeed, assume that the convex set A_ρ contains several corner points, as it is in Example 3. Each corner point admits more than one support hyperplane to A_ρ. Given

$\mathbf{M}$, any point $\rho(\mathbf{e})\mathbf{e}$ with $\mathbf{e} \in \operatorname{Arg} \min_{\mathbf{e} \in \mathbf{S}^{k-1}} \rho^2(\mathbf{e})\mathbf{e}'\mathbf{M}\mathbf{e}$ belongs both to ∂A_ρ and to the smallest ellipsoid of concentration $\{\mathbf{x} : \mathbf{x}'\mathbf{M}\mathbf{x} = c, c > 0\}$ still having common points with ∂A_ρ. Therefore, $\rho(\mathbf{e})\mathbf{e}$ is not a corner point, otherwise the preceding ellipsoid is not the smallest one.

From Theorems 1 and 2 a useful result follows helping in searching for the SS_ρ-optimal design.

Corollary 1. *Let the condition of Theorem 2 be fulfilled. If the* SS_ρ*-optimal design exists then it is also D-optimal and optimal with respect to the maximin criterion.*

In the next section we shall see that there exist situations when the reverse assertion of Corollary 1 also holds though it is not true in general.

3 Examples of the limit maximin criterion and $\mathrm{SS}_\rho(\varepsilon)$-optimal designs for the line fit model

Let us consider examples of the maximin criterion from Theorem 2 for the shape functions given in Section 2. It is easy to see that in all the cases the condition of Theorem 2 is fulfilled. We shall also supply the results with the $\mathrm{SS}_\rho(\varepsilon)$-optimal designs for the line fit model (2) on the interval $[0, 1]$ as well as on $[-1, 1]$.

Due to Lemma 1 from Liski and Zaigraev (2001), in the case $[0, 1]$ it is enough to take into account only designs of the form $\{0, 1; p, 1-p\}$, $0 < p < 1$ since they form a class of admissible designs. In the case $[-1, 1]$ due to the same argument it is enough to consider only designs of the form $\{-1, 1; p, 1-p\}$, $0 < p < 1$. That is why in both cases the $\mathrm{SS}_\rho(\varepsilon)$-optimal designs for any ρ are determined by the value p^* depending on ε, or more exactly on $\delta = \dfrac{\sqrt{n}\varepsilon}{\sigma}$.

Due to Lemma 2 from Liski and Zaigraev (2001), in Examples 1 and 3 the SS_ρ-optimal design on $[-1, 1]$ exists and is of the form $\{-1, 1; 0.5, 0.5\}$.

1. If $\rho(\mathbf{e}) \equiv 1$ then the $\mathrm{SS}_\rho(\varepsilon)$-criterion is simply the $\mathrm{DS}(\varepsilon)$-criterion. Here

$$\rho^2(\mathbf{e})\mathbf{e}'\mathbf{M}\mathbf{e} = \mathbf{e}'\mathbf{M}\mathbf{e}, \qquad \min_{\mathbf{e} \in S^{k-1}} \rho^2(\mathbf{e})\mathbf{e}'\mathbf{M}\mathbf{e} = \lambda_{\min}(\mathbf{M})$$

and the limit maximin criterion is the E-criterion.

For the line fit model (2) on the interval $[0, 1]$, the function $p^* = p^*(\delta)$ is monotonic and increases from $p_0 = 0.5$ (D-optimal design) to $p_\infty = 0.6$ (E-optimal design) along with increasing δ (see graph (1), Fig. 1). Clearly, there is no DS-optimal design here.

In general, the following characterization of design domination with respect to the DS-criterion for the case $k = 2$ takes place. Let $\lambda_1 \le \lambda_2$ be the eigenvalues of $\mathbf{M}(\xi_1)$ while $\mu_1 \le \mu_2$ be those of $\mathbf{M}(\xi_2)$.

Lemma 1. *Let* $k = 2$. *A design* ξ_1 *dominates* ξ_2 *with respect to the DS-criterion for the LSE of* $\boldsymbol{\beta}$ *if and only if*

$$\lambda_1 \ge \mu_1, \quad \lambda_1\lambda_2 \ge \mu_1\mu_2.$$

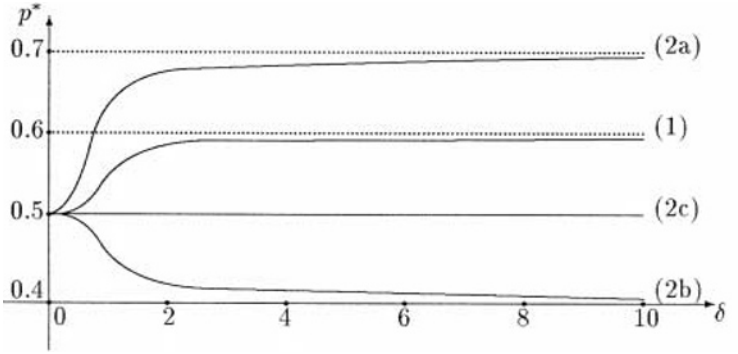

Fig. 1. The graph of $p^*(\delta)$ for (1), (2a), (2b), (2c).

Sufficiency follows from Marshall and Olkin (1979, Section 3.C, p. 64). Necessity follows from the limit behaviour of the DS(ε)-criterion. Indeed, let ξ_1 dominates ξ_2 with respect to the DS-criterion, that is with respect to the DS(ε)-criterion for all $\varepsilon > 0$. Using the limit behaviour of the DS(ε)-criterion, we get $\lambda_1 \geq \mu_1, \lambda_1\lambda_2 \geq \mu_1\mu_2$.

Corollary 2. *If* $k = 2$ *then a design is DS-optimal if and only if it is D-optimal and E-optimal.*

As we see, the existence of the DS-optimal design is determined by the behaviour of the DS(ε)-optimal designs when ε approaches 0 and ∞.

2. Let $\rho(\mathbf{e}) = (\mathbf{e}'\mathbf{D}\mathbf{e})^{-1/2}$, where $\mathbf{D}$ is a symmetric positive definite matrix. In this case

$$\rho^2(\mathbf{e})\mathbf{e}'\mathbf{M}\mathbf{e} = \frac{\mathbf{e}'\mathbf{M}\mathbf{e}}{\mathbf{e}'\mathbf{D}\mathbf{e}}, \quad \min_{\mathbf{e}\in S^{k-1}} \rho^2(\mathbf{e})\mathbf{e}'\mathbf{M}\mathbf{e} = \lambda_{\min}(\mathbf{D}^{-1}\mathbf{M})$$

due to the extremal properties of the above ratio (see e.g. Rao 1973, pp. 63, 74). Thus the maximin criterion is the criterion $\lambda_{\min}(\mathbf{D}^{-1}\mathbf{M}) \to \max$.

Consider the line fit model (2) on the interval $[0, 1]$ and denote

$$\mathbf{D}^{-1} = c\begin{pmatrix} d_1 & d_2 \\ d_2 & 1 \end{pmatrix}, \quad c > 0, d_1 > d_2^2.$$

It can be shown by direct calculation that the optimal design relative to the maximin criterion is determined by

$$p_\infty = \begin{cases} \dfrac{2d_1 + 3d_2 + 1}{4d_1 + 4d_2 + 1} & \text{if } d_1 + d_2 \geq 0 \\ 1 + d_2 & \text{if } d_1 + d_2 < 0. \end{cases} \tag{6}$$

It is interesting to note that p_∞ can take any value from the interval $(0, 1)$ depending on the choice of $\mathbf{D}$.

To show the possible dependence of p^* on δ for the $\mathrm{SS}_\rho(\varepsilon)$-optimal design, we consider three cases:

$$\text{(a) } \mathbf{D}^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad \text{(b) } \mathbf{D}^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}, \quad \text{(c) } \mathbf{D}^{-1} = \begin{pmatrix} 3 & -4 \\ -4 & 8 \end{pmatrix}.$$

It can be calculated from (6) that the corresponding values of p_∞ are:

$$\text{(a) } p_\infty = 0.7, \quad \text{(b) } p_\infty = 0.4, \quad \text{(c) } p_\infty = 0.5.$$

Relevant graphs (2a), (2b) and (2c) are given in Fig. 1. As one can see, they are quite different from each other. The first one looks like graph (1), except for the behaviour when $\delta \to \infty$. Here the function $p^*(\delta)$ increases monotonically from 0.5 to 0.7. But for (2b) the function $p^*(\delta)$ decreases monotonically from 0.5 to 0.4. At last, graph (2c) is a straight line $p^*(\delta) \equiv 0.5$. It means that there is no SS_ρ-optimal design in (a) and (b), but it exists in the case (c).

It turns out that it is not by occasion. There exists a class of matrices **D** for which the SS_ρ-optimal design exists. To find such a class, the next characterization result is helpful.

Lemma 2. *Let $k = 2$ and $\rho(\mathbf{e}) = (\mathbf{e}'\mathbf{D}\mathbf{e})^{-1/2}$, where* **D** *is a symmetric positive definite matrix. A design ξ_1 dominates ξ_2 with respect to the* SS_ρ*-criterion for the LSE of $\boldsymbol{\beta}$ if and only if*

$$\lambda_1 \geq \mu_1, \quad \lambda_1\lambda_2 \geq \mu_1\mu_2,$$

where $\lambda_1 \leq \lambda_2$ are the eigenvalues of $\mathbf{D}^{-1}\mathbf{M}(\xi_1)$, while $\mu_1 \leq \mu_2$ are those of $\mathbf{D}^{-1}\mathbf{M}(\xi_2)$.

The proof of Lemma 2 is similar to that of Lemma 1.

Again the existence of the SS_ρ-optimal design is determined by the behaviour of the $\mathrm{SS}_\rho(\varepsilon)$-optimal designs when ε approaches 0 and ∞. The SS_ρ-optimal design exists if and only if $p_0 = p_\infty = 0.5$ resulting in $p^*(\delta) \equiv 0.5$. From (6) it follows that $p_\infty = 0.5$ if and only if the matrix $\mathbf{D}^{-1}$ is of the form

$$\mathbf{D}^{-1} = c\begin{pmatrix} d & -0.5 \\ -0.5 & 1 \end{pmatrix}, \quad c > 0,\ d > 0.25.$$

Similar situation takes place also for the line fit model (2) on $[-1, 1]$. Here it can be shown by direct calculation that the optimal design relative to the maximin criterion is determined by

$$p_\infty = \frac{1}{2} + \frac{d_2}{2\max\{d_1, 1\}}.$$

Again p_∞ can take any value from the interval $(0, 1)$ depending on the choice of **D**. It is easy to see that $p_\infty = 0.5$ if and only if $d_2 = 0$. It means that the SS_ρ-optimal design exists if and only if the matrix **D** is diagonal.

3. Let $k = 2$ and $\rho(e_1, e_2) = \min\{a_1|e_1|^{-1}, a_2|e_2|^{-1}\}$, $a_1 > 0$, $a_2 > 0$. In this case $(e_1, e_2) = (\cos\phi, \sin\phi)$ and

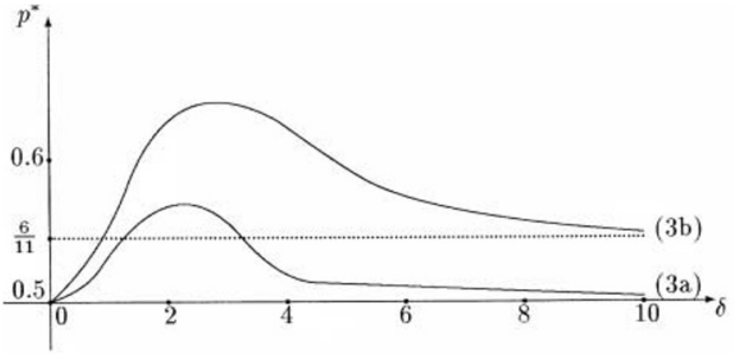

Fig. 2. The graph of $p^*(\delta)$ for (3a) and (3b).

$$\rho^2(\mathbf{e}(\phi))\mathbf{e}'(\phi)\mathbf{M}\mathbf{e}(\phi)$$
$$= \left(\min\left\{\frac{a_1}{|\cos\phi|}, \frac{a_2}{|\sin\phi|}\right\}\right)^2 (M_{11}\cos^2\phi + 2M_{12}\sin\phi\cos\phi + M_{22}\sin^2\phi),$$

where

$$\mathbf{M} = \begin{pmatrix} M_{11} & M_{12} \\ M_{12} & M_{22} \end{pmatrix}.$$

After some calculations we come to the following limit maximin criterion

$$\min_{\phi\in[0,2\pi]} \rho^2(\mathbf{e}(\phi))\mathbf{e}'(\phi)\mathbf{M}\mathbf{e}(\phi) = \frac{a_1^2 a_2^2 \det\mathbf{M}}{\max\{a_1^2 M_{11}, a_2^2 M_{22}\}} \to \max.$$

One can find by direct calculation that for the line fit model (2) on $[0, 1]$ the optimal design relative to the maximin criterion is determined by

$$p_\infty = \max\left\{\frac{1}{2}, 1 - \frac{a_1^2}{a_2^2}\right\}. \tag{7}$$

Hence, p_∞ can take any value from the interval $[0.5, 1)$ depending on the choice of a_1 and a_2.

To show the possible dependence of p^* on δ for the $\mathrm{SS}_p(\varepsilon)$-optimal design, we consider two cases:

$$\text{(a) } a_1 = a_2 = 1, \quad \text{(b) } a_1 = 1,\ a_2 = \sqrt{2.2}.$$

From (7) we get:

$$\text{(a) } p_\infty = \frac{1}{2}, \quad \text{(b) } p_\infty = \frac{6}{11}.$$

Relevant graphs (3a) and (3b) are given in Fig. 2.

In contrast to the previous cases, here the functions $p^*(\delta)$ are not monotonic. They are similar to each other, except for the limit behaviour as $\delta \to \infty$. As one can see, there is no SS_ρ-optimal designs here. Moreover, the situation is not determined by the behaviour of the $\mathrm{SS}_\rho(\varepsilon)$-optimal designs when ε approaches 0 and ∞. For example, in (3a) we have $p_0 = p_\infty = 0.5$ but $p^*(\delta) \neq 0.5$ for $0 < \delta < \infty$.

4 Concluding remarks

Dual problem. Dealing with probability $\mathrm{P}(\hat{\boldsymbol{\beta}}(\mathbf{M}) - \boldsymbol{\beta} \in \varepsilon A_\rho)$ for a given set A_ρ and given $\varepsilon > 0$, it is reasonable to make it as large as it is possible. But if the probability is given instead of ε, then it is reasonable to make ε as small as it is possible. Thus we come to the following dual problem: to choose an optimal design making ε the smallest possible for a given set A_ρ and given probability $\mathrm{P}(\hat{\boldsymbol{\beta}}(\mathbf{M}) - \boldsymbol{\beta} \in \varepsilon A_\rho) = \alpha \in (0, 1)$. In fact, $\hat{\boldsymbol{\beta}}(\mathbf{M}) - \varepsilon A_\rho$ looks like a *confidence set* of pregiven shape centered at $\hat{\boldsymbol{\beta}}(\mathbf{M})$ though it is not the case since this set depends on unknown σ.

We call a design α-*optimal* if the corresponding value of ε is the smallest one given α. Clearly, if there exists the SS_ρ-optimal design, then it is also α-optimal for any α. It is the case, e.g., for $\rho(\mathbf{e}) \equiv 1$ and the line fit model (2) on $[-1, 1]$. The α-optimal design is $\{-1, 1; 0.5, 0.5\}$ for any given α.

Consider the case when the SS_ρ-optimal design does not exist. If α is close to unit, then ε is necessarily sufficiently large. Then as we know from Theorem 2, the $\mathrm{SS}_\rho(\varepsilon)$-criterion is close to the maximin criterion. Therefore, the α-optimal design is close to the optimal design relative to the maximin criterion. For example, if $\rho(\mathbf{e}) \equiv 1$ and we consider the line fit model (2) on $[0, 1]$, then the α-optimal design is close to E-optimal. This closeness is higher along with increasing α.

Maximal probability content. The problems that have been touched in the paper can also be considered from another point of view. Namely, maximization of $\mathrm{P}(\hat{\boldsymbol{\beta}}(\mathbf{M}) - \boldsymbol{\beta} \in A)$ for all $A \in \mathscr{A}$ is a problem of searching for the maximum of the probability content simultaneously for all $A \in \mathscr{A}$ with respect to a given class of random vectors $\hat{\boldsymbol{\beta}}(\mathbf{M}) - \boldsymbol{\beta}$.

Problems of that type have been investigated by different authors. See Hall et al. (1980) and Mathew and Nordström (1997), for example. In those papers the probability content for the bivariate normally distributed random vector $\mathbf{Z} \sim \mathrm{N}_2(\mathbf{0}, \mathbf{I}_2)$ was considered and optimization was done with respect to a given class of sets that did not contain the origin, namely rotated squares and rotated ellipses. In our case taking in mind (5), k-variate normally distributed random vector $\mathbf{Z} \sim \mathrm{N}_k(\mathbf{0}, \mathbf{I}_k)$ is considered and maximization is done with respect to more complicated and general class of sets $\delta \mathbf{M}^{1/2} A_\rho$.

Integral criteria. Clearly, if $\mathscr{A}_1 \subset \mathscr{A}_2$ then a stochastic criterion of type (3) with $\mathscr{A}_2$ is stronger than that with $\mathscr{A}_1$. In other words, the richer $\mathscr{A}$ the stronger criterion (3).

In general, stochastic optimality criteria are quite strong, as we have seen. On the other hand, dealing with the shape criterion in practice it is more important to control the situation when ε is rather small or moderate while less important to know what happens when ε is sufficiently large.

Motivated by those arguments, one could consider the following two criteria both weaker than the DS-criterion:

$$\int_0^{\varepsilon} \mathrm{P}(\|\hat{\boldsymbol{\beta}}(\mathbf{M}) - \boldsymbol{\beta}\|^2 > t)\, dt \to \min \quad \forall \varepsilon > 0,$$

$$\int_{\varepsilon}^{\infty} \mathrm{P}(\|\hat{\boldsymbol{\beta}}(\mathbf{M}) - \boldsymbol{\beta}\|^2 > t)\, dt \to \min \quad \forall \varepsilon \geq 0.$$

In the paper by Mandal et al. (2000) the following weighted integral criterion was introduced called the *weighted coverage probability*:

$$\int_0^{\infty} \mathrm{P}(\|\hat{\boldsymbol{\beta}}(\mathbf{M}) - \boldsymbol{\beta}\|^2 \leq t) \frac{1}{2} e^{-t/2}\, dt \to \max.$$

Here the density function $\frac{1}{2}e^{-t/2}$ of a chi-squared distribution with 2 degrees of freedom plays a role of a weight function. Of course, any other reasonable weight function could be considered as well.

All those criteria will be investigated elsewhere.

Appendix

Proof of Theorem 2. We have due to (5),

$$\varphi_{\rho,\varepsilon}[\mathbf{M}] = \frac{\delta^k (\det \mathbf{M})^{1/2}}{(2\pi)^{k/2}} \int_{A_\rho} e^{-(\delta^2/2)\mathbf{z}'\mathbf{M}\mathbf{z}}\, d\mathbf{z} = \frac{a^{k/2} (\det \mathbf{M})^{1/2}}{\pi^{k/2}} \int_{A_\rho} e^{-a\mathbf{z}'\mathbf{M}\mathbf{z}}\, d\mathbf{z},$$

where $a = \dfrac{\delta^2}{2}$. Let us use spherical coordinates $(r, \phi_1, \ldots, \phi_{k-1})$:

$$z_1 = r \cos\phi_1 \cos\phi_2 \ldots \cos\phi_{k-1} = re_1(\boldsymbol{\phi}),$$

$$z_2 = r \sin\phi_1 \cos\phi_2 \ldots \cos\phi_{k-1} = re_2(\boldsymbol{\phi}),$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$z_{k-1} = r \sin\phi_{k-2} \cos\phi_{k-1} = re_{k-1}(\boldsymbol{\phi}),$$

$$z_k = r \sin\phi_{k-1} = re_k(\boldsymbol{\phi}),$$

with Jacobian $r^{k-1}J(\boldsymbol{\phi})$, $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_{k-1}) \in \boldsymbol{\Phi} = [0, 2\pi] \times \cdots \times [0, 2\pi] \times \left[-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right]$, where $J(\boldsymbol{\phi}) = |\prod_{i=1}^{k-1} \cos^{i-1}\phi_i|$. Denote $\mathbf{e}(\boldsymbol{\phi}) = (e_1(\boldsymbol{\phi}), \ldots, e_k(\boldsymbol{\phi}))'$.

Applying them and changing variable $r = \rho(\mathbf{e}(\boldsymbol{\phi}))z$, we obtain

$$1 - \varphi_{\rho,\varepsilon}[\mathbf{M}] = \frac{a^{k/2}(\det \mathbf{M})^{1/2}}{\pi^{k/2}} \int_{\Phi} J(\boldsymbol{\phi}) \left(\int_{\rho(\mathbf{e}(\boldsymbol{\phi}))}^{\infty} r^{k-1} e^{-ar^2 \mathbf{e}(\boldsymbol{\phi})'\mathbf{M}\mathbf{e}(\boldsymbol{\phi})} \, dr \right) d\boldsymbol{\phi}$$

$$= \frac{a^{k/2}(\det \mathbf{M})^{1/2}}{\pi^{k/2}} \int_{\Phi} \rho^k(\mathbf{e}(\boldsymbol{\phi})) J(\boldsymbol{\phi}) I(a\rho^2(\mathbf{e}(\boldsymbol{\phi}))\mathbf{e}(\boldsymbol{\phi})'\mathbf{M}\mathbf{e}(\boldsymbol{\phi})) \, d\boldsymbol{\phi},$$

where

$$I(x) = \int_1^{\infty} z^{k-1} e^{-xz^2} \, dz = x^{-k/2} \int_{x^{1/2}}^{\infty} u^{k-1} e^{-u^2} \, du.$$

Due to L'Hospital's rule, $I(x) \sim \dfrac{e^{-x}}{2x}$ as $x \to \infty$. Therefore, we get as $a \to \infty$,

$$1 - \varphi_{\rho,\varepsilon}[\mathbf{M}] \sim \frac{a^{k/2-1}(\det \mathbf{M})^{1/2}}{2\pi^{k/2}} \int_{\Phi} \frac{\rho^{k-2}(\mathbf{e}(\boldsymbol{\phi})) J(\boldsymbol{\phi}) e^{-a\rho^2(\mathbf{e}(\boldsymbol{\phi}))\mathbf{e}(\boldsymbol{\phi})'\mathbf{M}\mathbf{e}(\boldsymbol{\phi})}}{\mathbf{e}(\boldsymbol{\phi})'\mathbf{M}\mathbf{e}(\boldsymbol{\phi})} d\boldsymbol{\phi}.$$

Our goal is to establish the asymptotics of the integral

$$\int_{\Phi} c(\mathbf{e}(\boldsymbol{\phi})) e^{-ah(\mathbf{e}(\boldsymbol{\phi}))} \, d\boldsymbol{\phi} \tag{8}$$

when $a \to \infty$, where

$$c(\mathbf{e}(\boldsymbol{\phi})) = \frac{\rho^{k-2}(\mathbf{e}(\boldsymbol{\phi})) J(\boldsymbol{\phi})}{\mathbf{e}'(\boldsymbol{\phi})\mathbf{M}\mathbf{e}(\boldsymbol{\phi})}, \quad h(\mathbf{e}(\boldsymbol{\phi})) = \rho^2(\mathbf{e}(\boldsymbol{\phi}))\mathbf{e}'(\boldsymbol{\phi})\mathbf{M}\mathbf{e}(\boldsymbol{\phi}).$$

Given $\mathbf{M}$ define

$$\Phi^* = \left\{ \boldsymbol{\phi} \in \Phi : \mathbf{e}(\boldsymbol{\phi}) \in \operatorname{Arg} \min_{\mathbf{e} \in \mathbf{S}^{k-1}} \rho^2(\mathbf{e})\mathbf{e}'\mathbf{M}\mathbf{e} \right\}.$$

We shall thoroughly consider the situation when Φ^* consists of a single point $\boldsymbol{\phi}^*$ and show that only a neighbourhood of $\boldsymbol{\phi}^*$ is essential in establishing the asymptotics of integral (8) as $a \to \infty$. It is worth noting that without loss of generality we can always assume that $J(\boldsymbol{\phi}^*) \neq 0$.

Denote by h^* the point of minimum for the function $h(\mathbf{e}(\boldsymbol{\phi}))$, that is $h^* = \rho^2(\mathbf{e}(\boldsymbol{\phi}^*))\mathbf{e}(\boldsymbol{\phi}^*)'\mathbf{M}\mathbf{e}(\boldsymbol{\phi}^*)$. Under the condition of the theorem, the function $h(\mathbf{e}(\boldsymbol{\phi}))$ in a neighbourhood of $\boldsymbol{\phi}^*$ admits the Taylor expansion

$$h(\mathbf{e}(\boldsymbol{\phi})) = h^* + \frac{1}{2}(\boldsymbol{\phi} - \boldsymbol{\phi}^*)'\mathbf{H}(\tilde{\boldsymbol{\phi}})(\boldsymbol{\phi} - \boldsymbol{\phi}^*), \tag{9}$$

where $\mathbf{H}(\tilde{\boldsymbol{\phi}})$ is the *hessian* of $h(\mathbf{e}(\boldsymbol{\phi}))$ at $\tilde{\boldsymbol{\phi}} = \boldsymbol{\phi}^* + \theta(\boldsymbol{\phi} - \boldsymbol{\phi}^*)$ for some θ, $|\theta| \leq 1$.

Since ρ is continuous, for any sufficiently small $\gamma > 0$ we have

$$\min_{\boldsymbol{\phi} \in \Phi_1} h(\mathbf{e}(\boldsymbol{\phi})) = h^* + c(\gamma),$$

where $c(\gamma) > 0$ and

$$\Phi_1 = \{\boldsymbol{\phi} \in \Phi : \|\boldsymbol{\phi} - \boldsymbol{\phi}^*\| \geq \gamma\}.$$

Now we partition the integral in (8) onto two the parts, namely

$$\int_{\Phi} c(\mathbf{e}(\boldsymbol{\phi}))e^{-ah(\mathbf{e}(\boldsymbol{\phi}))}\,d\boldsymbol{\phi} = \int_{\Phi_1} c(\mathbf{e}(\boldsymbol{\phi}))e^{-ah(\mathbf{e}(\boldsymbol{\phi}))}\,d\boldsymbol{\phi} + \int_{\Phi_2} c(\mathbf{e}(\boldsymbol{\phi}))e^{-ah(\mathbf{e}(\boldsymbol{\phi}))}\,d\boldsymbol{\phi} = I_1 + I_2, \tag{10}$$

where

$$\Phi_2 = \{\boldsymbol{\phi} \in \Phi : \|\boldsymbol{\phi} - \boldsymbol{\phi}^*\| < \gamma\}.$$

The first integral can be estimated as

$$I_1 \leq e^{-a(h^*+c(\gamma))} \int_{\Phi_1} c(\mathbf{e}(\boldsymbol{\phi}))\,d\boldsymbol{\phi}$$

and in particular,

$$I_1 = o(a^{-(k-1)/2}e^{-ah^*}) \quad \text{as } a \to \infty. \tag{11}$$

Under the condition of the theorem for sufficiently small $\gamma > 0$ there exists such $0 < d(\gamma) < 1$ that for all $\boldsymbol{\phi} \in \Phi_2$ we have

$$|c(\mathbf{e}(\boldsymbol{\phi})) - c(\mathbf{e}(\boldsymbol{\phi}^*))| < d(\gamma)c(\mathbf{e}(\boldsymbol{\phi}^*)), \tag{12}$$

$$|\mathbf{e}'\mathbf{H}(\tilde{\boldsymbol{\phi}})\mathbf{e} - \mathbf{e}'\mathbf{H}(\boldsymbol{\phi}^*)\mathbf{e}| < d(\gamma)\mathbf{e}'\mathbf{H}(\boldsymbol{\phi}^*)\mathbf{e} \quad \forall \mathbf{e} \in S^{k-2}. \tag{13}$$

Thus from (9), (12) and (13) we get

$$(1 - d(\gamma))c(\mathbf{e}(\boldsymbol{\phi}^*))e^{-ah^*} \int_{\Phi_2} e^{-(a/2)(1+d(\gamma))(\boldsymbol{\phi}-\boldsymbol{\phi}^*)'\mathbf{H}(\boldsymbol{\phi}^*)(\boldsymbol{\phi}-\boldsymbol{\phi}^*)}\,d\boldsymbol{\phi} < I_2$$
$$< (1 + d(\gamma))c(\mathbf{e}(\boldsymbol{\phi}^*))e^{-ah^*} \int_{\Phi_2} e^{-(a/2)(1-d(\gamma))(\boldsymbol{\phi}-\boldsymbol{\phi}^*)'\mathbf{H}(\boldsymbol{\phi}^*)(\boldsymbol{\phi}-\boldsymbol{\phi}^*)}\,d\boldsymbol{\phi},$$

or equivalently

$$(1 - d(\gamma))a^{(k-1)/2} \int_{\{\mathbf{u}:\mathbf{u}+\boldsymbol{\phi}^* \in \Phi_2,\, \|\mathbf{u}\|<\gamma\}} e^{-(a/2)(1+d(\gamma))\mathbf{u}'\mathbf{H}(\boldsymbol{\phi}^*)\mathbf{u}}\,d\mathbf{u} < \frac{a^{(k-1)/2}e^{ah^*}I_2}{c(\mathbf{e}(\boldsymbol{\phi}^*))}$$
$$< (1 + d(\gamma))a^{(k-1)/2} \int_{\{\mathbf{u}:\mathbf{u}+\boldsymbol{\phi}^* \in \Phi_2,\, \|\mathbf{u}\|<\gamma\}} e^{-(a/2)(1-d(\gamma))\mathbf{u}'\mathbf{H}(\boldsymbol{\phi}^*)\mathbf{u}}\,d\mathbf{u},$$

that is

$$\frac{1-d(\gamma)}{(1+d(\gamma))^{(k-1)/2}}\int_{\{\mathbf{v}:\mathbf{v}/(\sqrt{a}\sqrt{1+d(\gamma)})+\boldsymbol{\phi}^*\in\Phi_2,\,\|\mathbf{v}\|<\gamma\sqrt{a}\sqrt{1+d(\gamma)}\}} e^{-(1/2)\mathbf{v}'\mathbf{H}(\boldsymbol{\phi}^*)\mathbf{v}}\,d\mathbf{v}$$
$$<\frac{a^{(k-1)/2}e^{ah^*}I_2}{c(\mathbf{e}(\boldsymbol{\phi}^*))}$$
$$<\frac{1+d(\gamma)}{(1-d(\gamma))^{(k-1)/2}}\int_{\{\mathbf{v}:\mathbf{v}/(\sqrt{a}\sqrt{1+d(\gamma)})+\boldsymbol{\phi}^*\in\Phi_2,\,\|\mathbf{v}\|<\gamma\sqrt{a}\sqrt{1-d(\gamma)}\}} e^{-(1/2)\mathbf{v}'\mathbf{H}(\boldsymbol{\phi}^*)\mathbf{v}}\,d\mathbf{v}.$$

Finally, when a approaches ∞ we obtain

$$\frac{(1-d(\gamma))(2\pi)^{(k-1)/2}}{(1+d(\gamma))^{(k-1)/2}(\det\mathbf{H}(\boldsymbol{\phi}^*))^{1/2}}<\liminf_{a\to\infty}\frac{a^{(k-1)/2}e^{ah^*}I_2}{c(\mathbf{e}(\boldsymbol{\phi}^*))}$$
$$\le\limsup_{a\to\infty}\frac{a^{(k-1)/2}e^{ah^*}I_2}{c(\mathbf{e}(\boldsymbol{\phi}^*))}<\frac{(1+d(\gamma))(2\pi)^{(k-1)/2}}{(1-d(\gamma))^{(k-1)/2}(\det\mathbf{H}(\boldsymbol{\phi}^*))^{1/2}},$$

or

$$\frac{a^{(k-1)/2}e^{ah^*}I_2}{c(\mathbf{e}(\boldsymbol{\phi}^*))}\sim\frac{(2\pi)^{(k-1)/2}}{(\det\mathbf{H}(\boldsymbol{\phi}^*))^{1/2}}\quad\text{as } a\to\infty \tag{14}$$

since $d(\gamma)$ can be arbitrary small along with γ. From (10), (11) and (14) it follows that as $a\to\infty$,

$$\int_{\Phi}c(\mathbf{e}(\boldsymbol{\phi}))e^{-ah(\mathbf{e}(\boldsymbol{\phi}))}\,d\boldsymbol{\phi}\sim\frac{(2\pi)^{(k-1)/2}c(\mathbf{e}(\boldsymbol{\phi}^*))e^{-ah^*}}{a^{(k-1)/2}(\det\mathbf{H}(\boldsymbol{\phi}^*))^{1/2}}$$

and therefore

$$1-\varphi_{\rho,\varepsilon}[\mathbf{M}]\sim\frac{2^{(k-3)/2}(\det\mathbf{M})^{1/2}c(\mathbf{e}(\boldsymbol{\phi}^*))e^{-ah^*}}{(a\pi)^{1/2}(\det\mathbf{H}(\boldsymbol{\phi}^*))^{1/2}}. \tag{15}$$

Now it is easy to see that the same matrix $\mathbf{M}^*$ that maximizes $h^*=\min_{\boldsymbol{\phi}\in\Phi}\rho^2(\mathbf{e}(\boldsymbol{\phi}))\mathbf{e}(\boldsymbol{\phi})'\mathbf{M}\mathbf{e}(\boldsymbol{\phi})$ also maximizes $\varphi_{\rho,\varepsilon}[\mathbf{M}]$ for sufficiently large $a=\dfrac{n\varepsilon^2}{2\sigma^2}$ and vice versa.

If Φ^* consists of $m>1$ points $\boldsymbol{\phi}_1^*,\ldots,\boldsymbol{\phi}_m^*$ then we partition integral (8) onto the sum of m integrals, each over a subset consisting only of a single point from Φ^*. Similarly, instead of (15) we obtain

$$1-\varphi_{\rho,\varepsilon}[\mathbf{M}]\sim\frac{2^{(k-3)/2}(\det\mathbf{M})^{1/2}e^{-ah^*}}{(a\pi)^{1/2}}\sum_{i=1}^{m}\frac{c(\mathbf{e}(\boldsymbol{\phi}_i^*))}{(\det\mathbf{H}(\boldsymbol{\phi}_i^*))^{1/2}}$$

and one can make the same conclusion as above.

The variety of possible sets Φ^* containing infinitely many points is extremely rich. But in each case in the same way as above it can be shown that the asymptotic expression for $1 - \varphi_{\rho,\varepsilon}[\mathbf{M}]$ has a term e^{-ah^*} that plays a crucial role in minimizing $1 - \varphi_{\rho,\varepsilon}[\mathbf{M}]$ for sufficiently large ε. For example, if Φ^* is a connected set of dimension $k - 1$ then we partition integral (8) onto the sum of three integrals: the first one over Φ^*, the second one over a γ-neighbourhood of Φ^* and the rest. The last integral is asymptotically as $a \to \infty$ unessential while the first two integrals asymptotically both have a term e^{-ah^*}.

Acknowledgement. The author would like to express his gratitude to the referees for careful reading of the manuscript and valuable remarks which essentially improved an earlier version of the paper.

References

Eaton ML, Perlman MD (1991) Concentration inequalities for multivariate distributions: I. Multivariate normal distributions. Statistics and Probability Letters 12:487–504

Giovagnoli A, Wynn HP (1995) Multivariate dispersion orderings. Statistics and Probability Letters 22:325–332

Hall RL, Kanter M, Perlman MD (1980) Inequalities for the probability content of a rotated square and related convolutions. Annals of Probability 8(4):802–813

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

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

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

Liski E, Zaigraev A (2001) A stochastic characterization of Loewner optimality design criterion in linear models. Metrika 53:207–222

Mandal NK, Shah KR, Sinha BK (2000) Comparison of test vs. control treatments using distance optimality criterion. Metrika 52:147–162

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

Mathew T, Nordström K (1997) Inequalities for the probability content of a rotated ellipse and related stochastic domination results. Annals of Applied Probability 7(4):1106–1117

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

Rao CR (1973) Linear statistical inference and its applications, 2nd edition. Wiley, New York

Shah KR, Sinha BK (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 BK (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