Acta Math. Univ. Comenianae
Vol. LXXVII, 2(2008), pp. 155–166

COMPUTING THE MINIMAL EFFICIENCY OF DESIGNS BY A DIFFERENTIABLE APPROXIMATION OF Φ_{E_k}-OPTIMALITY CRITERIA

L. BUŠOVÁ

Abstract. Consider the linear regression model with uncorrelated errors and an experimental design ξ. In the paper, we propose a numerical method for calculating the minimal efficiency of ξ in the class $\mathbb{O}$ of orthogonally invariant information criteria. For this purpose, we introduce the concept of $\Phi_{k,p}^{(m)}$-optimality criteria. Then we show that $\Phi_{E_k}^{(m)}$ criteria can be differentiably approximated by $\Phi_{k,p}^{(m)}$ criteria, therefore it allows us to use standard numerical procedures to arrive at boundaries for $\Phi_{E_k}^{(m)}$ optimal values, and hence at the intervals for the minimal efficiency of designs under the class of all orthogonally invariant information criteria. The approach is then illustrated on the polynomial model of degrees $2, \dots, 8$.

1. Introduction

The aim of this article is to numerically calculate boundaries for the minimal efficiency of designs with respect to the class of orthogonally invariant information criteria $\mathbb{O}$, containing many well-known criteria, e.g. all Kiefer's Φ_p-optimality criteria (see e.g. [**8**, p. 94] or [**9**, p. 139]). For this purpose, it turns out to be useful to study partial sums of eigenvalues of information matrices. As an example, the sums of k smallest (or largest) eigenvalues can be used to characterize universal optimality in the class $\mathbb{O}$, as was shown in the article of Bondar [**1**]. The sums of k smallest eigenvalues, viewed as special orthogonally invariant information functions $\Phi_{E_k}^{(m)}$, are studied in [**3**]. The main result of the paper [**3**] is that minimal efficiency of a given design with respect to $\mathbb{O}$ can be calculated using the optimal values of $\Phi_{E_k}^{(m)}$-criteria. In some simpler models, the optimal values can be calculated exactly, with the help of equivalence theorem for criteria of $\Phi_{E_k}^{(m)}$-optimality based on special results from convex analysis (see [**7**]). Moreover, boundaries for the optimal values of $\Phi_{E_k}^{(m)}$-criteria can be calculated from Φ_p-optimal designs (see [**4**]), but these boundaries are mostly very conservative. In general, calculation of the $\Phi_{E_k}^{(m)}$-optimal values and the minimal efficiency of designs is difficult, mainly because the criteria are nondifferentiable, and the standard routines for calculating the optimal designs cannot be applied. In our article, we solve this problem

Received May 9, 2005; revised February 22, 2008.
2000 *Mathematics Subject Classification.* Primary 62K05.

numerically by replacing $\Phi_{E_k}^{(m)}$-criteria with suitable differentiable approximations.

Consider the linear regression model

$$y(x) = f^T(x)\theta + \varepsilon \tag{1}$$

on a compact experimental domain $\mathcal{X} \subseteq \mathbb{R}^n$, where $f : \mathcal{X} \to \mathbb{R}^m$ is a vector of known continuous regression functions, $\theta \in \mathbb{R}^m$ is a vector of unknown parameters, and ε is a random error. The errors are assumed to be uncorrelated for different observations.

By an experimental design we understand a probability measure ξ on $\mathcal{X}$ with a finite support $\mathcal{X}_\xi = \{x \in \mathcal{X},\ \xi(x) > 0\}$. By Ξ we denote the set of all designs on $\mathcal{X}$. The value $\xi(x)$ is understood as the relative frequency of replications to be taken in x. To be able to determine the quality of the design, we need to define an optimality criterion measuring the amount of information about parameters that can be gained from the experiment based on the design. As it is usual in the optimal design literature, we will focus our attention on the information which is contained in the positive semidefinite information matrix defined as

$$M(\xi) = \sum_{x \in \mathcal{X}_\xi} f(x) f^T(x) \xi(x)$$

for $\xi \in \Xi$.

Optimality criterion Φ is a real-valued function defined on the set $\mathcal{S}_+^m$ of all positive semidefinite matrices of type $m \times m$. A design ξ^* is Φ-optimal iff $\xi^* = \arg\max_{\xi \in \Xi} \Phi(M(\xi))$, and $M(\xi^*)$ is called a Φ-optimal information matrix. Through the article, we will use the following special type of optimality criteria.

Function $\Phi : \mathcal{S}_+^m \to \langle 0, \infty)$ is called an information function ([**9**]), if it is not identically zero and if it satisfies the following conditions:

- isotonicity:

$$D - C \in \mathcal{S}_+^m \Rightarrow \Phi(C) \leq \Phi(D) \qquad \forall C, D \in \mathcal{S}_+^m$$

- concavity:

$$\Phi(\alpha C + (1-\alpha) D) \geq \alpha \Phi(C) + (1-\alpha)\Phi(D) \qquad \forall C, D \in \mathcal{S}_+^m,\ \alpha \in (0,1)$$

- positive homogeneity:

$$\Phi(\alpha C) = \alpha \Phi(C) \qquad \forall C \in \mathcal{S}_+^m,\ \alpha \geq 0$$

- upper semicontinuity: the sets $\{C \in \mathcal{S}_+^m; \Phi(C) \geq c\}$ are closed for all $c \in \mathbb{R}$

In this class we can find almost all commonly used criteria in their concave and positive homogeneous versions.

If Φ is an information function, then a Φ-optimal design ξ^* always exists and $\Phi(M(\xi^*)) > 0$ ([**9**, p. 117]).

From the practical point of view, interesting characteristics of a design is its quality compared to the Φ-optimal design. The measure of this quality is called the Φ-efficiency of the design ξ.

The Φ-efficiency of the design ξ is defined as follows ([**9**, p. 115]):

$$\text{eff}(M(\xi) \mid \Phi) = \frac{\Phi(M(\xi))}{\max_{\zeta \in \Xi} \Phi(M(\zeta))}. \tag{2}$$

Our aim in the article will be to propose a numerical method of computing minimal efficiency of a design ξ with respect to the wide class of all orthogonally invariant information criteria which will be defined in the next section.

2. The class of orthogonally invariant information criteria

Definition 2.1. We define the class $\mathbb{O}$ of all orthogonally invariant information criteria as the set of information functions $\Phi : \mathcal{S}_+^m \to \langle 0, \infty)$ which satisfy the condition of orthogonal invariance, i.e. $\Phi(UCU^T) = \Phi(C)$ for all $C \in \mathcal{S}_+^m$ and orthogonal matrices U of type $m \times m$.

This class encompasses many well-known optimality criteria, e.g. Kiefer's Φ_p criteria with their widely used special cases: D, A, and E criteria, as well as their convex combinations.

Definition 2.2. For m-dimensional model and $p \in \langle -\infty, 1\rangle$ the Kiefer's criterion $\Phi_p^{(m)} : \mathcal{S}_+^m \to \langle 0, \infty)$ is defined as follows:

$$\Phi_p^{(m)}(M) = \begin{cases} \left(\frac{1}{m}\sum_{i=1}^m \lambda_i^p(M)\right)^{1/p} & \text{if } p \in (-\infty, 0) \text{ and M is regular or if } p \in (0, 1\rangle \\ \left(\prod_{i=1}^m \lambda_i(M)\right)^{1/m} & \text{if } p = 0 \\ \lambda_1(M) & \text{if } p = -\infty \\ 0 & \text{if } p \in (-\infty, 0) \text{ and M is singular} \end{cases}$$

where $\lambda(A) = (\lambda_1(A), \lambda_2(A), \ldots, \lambda_m(A))^T$ is the vector of all eigenvalues of A in nondecreasing order defined on the set $\mathcal{S}_+^m$, i.e.

$$0 \le \lambda_1(A) \le \lambda_2(A) \le \ldots \le \lambda_m(A).$$

Note that Φ is orthogonally invariant iff its value depends only on the eigenvalues of the matrix, i.e. iff $\lambda(M_1) = \lambda(M_2) \Rightarrow \Phi(M_1) = \Phi(M_2)$.

For finding the minimal efficiency of a design in the set of all orthogonally invariant information criteria we will use the $\Phi_{E_k}^{(m)}$-optimality criteria, introduced in [**3**].

Definition 2.3. $\Phi_{E_k}^{(m)} : \mathcal{S}_+^m \to \langle 0, \infty)$ is an optimality criterion defined as the sum of k smallest eigenvalues of the information matrix:

$$\Phi_{E_k}^{(m)}(M) = \sum_{i=1}^k \lambda_i(M). \tag{3}$$

Theorem 2.1. *The minimal efficiency theorem ([3]). Let $\xi \in \Xi$. Then*

$$\inf_{\Phi \in \mathbb{O}} \operatorname{eff}(M(\xi) \mid \Phi) = \min_{k=1,\dots,m} \operatorname{eff}(M(\xi) \mid \Phi_{E_k}^{(m)}). \tag{4}$$

A consequence of the theorem is that for finding the minimal efficiency of the design ξ on the whole class $\mathbb{O}$ we only need to find the minimal efficiencies for $\Phi_{E_k}^{(m)}$, $k = 1, \dots, m$. This can be difficult because for $k < m$ the criteria $\Phi_{E_k}^{(m)}$ are not differentiable everywhere and standard numerical procedures can be impossible to apply.

In the article, we solve the problem by implementing a special class of criteria, which will be used as a differentiable approximation of the $\Phi_{E_k}^{(m)}$ criteria.

3. $\Phi_{k,p}^{(m)}$-optimality criteria

Definition 3.1. For m-dimensional model, integer $k \in \{1, \dots, m\}$, and $p \in \langle -\infty, 0\rangle$ we define the criterion $\Phi_{k,p}^{(m)} : \mathcal{S}_+^m \to \langle 0, \infty)$ as follows:
if $\operatorname{rank}(M) \geq m - k + 1$,

$$\Phi_{k,p}^{(m)}(M) = \begin{cases} \left(\binom{m}{k}^{-1} \sum_{1 \leq i_1 < \dots < i_k \leq m} \left(\sum_{j=1}^{k} \lambda_{i_j}(M) \right)^p \right)^{1/p} & \text{if } p \in (-\infty, 0) \\ \left(\prod_{1 \leq i_1 < \dots < i_k \leq m} \sum_{j=1}^{k} \lambda_{i_j}(M) \right)^{\binom{m}{k}^{-1}} & \text{if } p = 0 \\ \sum_{j=1}^{k} \lambda_j(M) & \text{if } p = -\infty \end{cases}$$

if $\operatorname{rank}(M) < m - k + 1$, then $\Phi_{k,p}^{(m)}(M) = 0$

Substituting for $k = 1$ we get Kiefer's $\Phi_p^{(m)}$ criteria and for $p = -\infty$ we get $\Phi_{E_k}^{(m)}$ criteria.

3.1. Properties of $\Phi_{k,p}^{(m)}$ criteria

In this subsection, we will prove that $\Phi_{k,p}^{(m)}$ criteria belong to the class $\mathbb{O}$ and we will derive the gradient of $\Phi_{k,p}^{(m)}$.

Definition 3.2. For $p \in \langle -\infty, 0\rangle$, $s \in \mathbb{N}$ we define a function $\varphi_p^{(s)} : \langle 0,\infty)^s \to \langle 0,\infty)$:

$$\varphi_p^{(s)}(\lambda_1,\dots,\lambda_s) = \begin{cases} \left(\frac{1}{s}\sum_{i=1}^{s}\lambda_i^p\right)^{1/p} & \text{if } p \in (-\infty,0) \text{ and } \lambda_i > 0\ \forall i \\ \left(\prod_{i=1}^{s}\lambda_i\right)^{1/s} & \text{if } p = 0 \\ \min(\lambda_1,\dots,\lambda_s) & \text{if } p = -\infty \\ 0 & \text{if } p \in (-\infty,0) \text{ and } \lambda_i = 0 \text{ for some i} \end{cases}$$

Lemma 3.1. *The function $\varphi_p^{(s)}$ is continuous and concave on $\langle 0,\infty)^s$.*

Proof. To prove continuity of $\varphi_p^{(s)}$, we will use the continuity of $\Phi_p^{(s)}$ criteria, which is a well-known fact in optimal design. The function $\mathrm{diag}(\cdot) : \langle 0,\infty)^s \to \mathcal{S}_+^s$, mapping the vector a onto a diagonal matrix of type $s \times s$ with diagonal elements equal to the components of a, is continuous on $\langle 0,\infty)^s$. Because $\varphi_p^{(s)}(\lambda) = \Phi_p^{(s)}(\mathrm{diag}(\lambda))\ \forall\lambda \in \langle 0,\infty)^s$, $\varphi_p^{(s)}$ is composed of two continuous functions, i.e $\varphi_p^{(s)}$ is continuous itself.

Next, we show that $\varphi_p^{(s)} : \langle 0,\infty)^s \to \langle 0,\infty)$ is concave. Let $\Psi_p^{(s)}(M) = \Phi_p^{(s)}(M)$ for $M \in \mathcal{S}_+^s$ and $\Psi_p^{(s)}(M) = -\infty$ for $M \in \mathcal{S}^s \backslash \mathcal{S}_+^s$, where $\mathcal{S}^s$ is the set of all symmetric matrices of type $s \times s$. Using the concavity of $\Phi_p^{(s)}$, it is easy to show that $\Psi_p^{(s)}$ is concave on the linear space $\mathcal{S}^s$. Moreover, the function $\mathrm{diag} : \mathbb{R}^s \to \mathcal{S}^s$ is linear. Hence the composition $\Psi_p^{(s)} \circ \mathrm{diag} : \mathbb{R}^s \to \mathbb{R} \cup \{-\infty\}$ is a concave function ([**11**, p. 38]) (the linear space $\mathcal{S}_+^s$ can be considered as $\mathbb{R}^{(s^2+s)/2}$). But $\varphi_p^{(s)}$ is a restriction of $\Psi_p^{(s)} \circ \mathrm{diag}$ on the set $\langle 0,\infty)^s$, therefore $\varphi_p^{(s)}$ is concave itself. □

Lemma 3.2.

$$\Phi_{k,p}^{(m)}(M) = \varphi_p^{\binom{m}{k}}(L_k^{(m)}(\lambda(M))), \qquad \forall M \in \mathcal{S}_+^m$$

where $L_k^{(m)} : \mathbb{R}^m \to \mathbb{R}^{\binom{m}{k}}$ is a linear function such that

$$L_k^{(m)} : (\lambda_1,\dots,\lambda_m) \to (\sum_{j\in I_1}\lambda_j,\dots,\sum_{j\in I_{\binom{m}{k}}}\lambda_j)$$

$I_1,\dots,I_{\binom{m}{k}}$ *are all subsets of $\{1,\dots,m\}$ with k elements.*

Theorem 3.1. *The optimality criterion $\Phi_{k,p}^{(m)}$ belongs to the class of orthogonally invariant information criteria.*

Proof.

1. Isotonicity and orthogonal invariance follow from [**3**, Proposition 1].
2. Positive homogeneity: it is obvious that for every $M \in \mathcal{S}_+^m$ and for every $\alpha > 0$: $\lambda(\alpha M) = \alpha\lambda(M)$. Hence positive homogeneity follows directly from the definition of $\Phi_{k,p}^{(m)}$ criteria.

3. Continuity: the function $\lambda : \mathcal{S}_+^m \to \mathbb{R}_+^m$ is continuous ([**6**, p. 540]). The function $\varphi_p^{\binom{m}{k}}$ is continuous (Lemma 3.1), as well as the function $L_k^{(m)}$, because it is linear. As it is evident from Lemma 3.2, $\Phi_{k,p}^{(m)}$ is composed of continuous functions $\varphi_p^{\binom{m}{k}}$, $L_k^{(m)}$ and λ which means that $\Phi_{k,p}^{(m)}$ is continuous itself.
4. Concavity: $\varphi_p^{\binom{m}{k}}$ is concave, therefore $\varphi_p^{\binom{m}{k}} \circ L_k^{(m)}$ is concave. Using the Davis theorem [**2**] and Lemma 3.2, it follows that $\Phi_{k,p}^{(m)}$ is concave.

□

Now we will use [**2**] to derive a formula for the gradient of the function $\Phi_{k,p}^{(m)}$. Note that by a symmetric function we mean a function $f : \mathbb{R}^n \to \mathbb{R}^n$ with a property $f(x) = f(y)$, if the vector $y \in \mathbb{R}^n$ is a permutation of the vector $x \in \mathbb{R}^n$.

Lemma 3.3. [**2**, p. 105] *Let $f : \langle 0, \infty)^s \to \langle 0, \infty)$ be symmetric, concave, and differentiable function on $(0, \infty)^s$ and let $\Phi : \mathcal{S}_{++}^s \to \langle 0, \infty)$ be defined as $\Phi = f \circ \lambda$. Then Φ is differentiable on $\mathcal{S}_{++}^s$ and:*

$$\bigtriangledown\Phi(M) = U \operatorname{diag}(\bigtriangledown f(\lambda(M)))U^T \tag{5}$$

where U is a matrix, for which $M = U \operatorname{diag}(\lambda(M))U^T$ and $\bigtriangledown f$ denotes the gradient of f.

Theorem 3.2. *Let $k \in \{1, \ldots, m\}$, $p \in (-\infty, 0)$, $\operatorname{rank}(M) \geq m-k+1$. Denote for $1 \leq i \leq m$:*

$$\delta_{i,k}(\lambda) = \sum_{\substack{1 \leq i_1 < \ldots < i_k \leq m \\ i \in \{i_1, \ldots, i_k\}}} \left(\sum_{j=1}^{k} \lambda_{i_j} \right)^{p-1}, \qquad \lambda \in (0, \infty)^m.$$

Then the gradient of the function $\Phi_{k,p}^{(m)}$ is:

$$\begin{aligned} &\bigtriangledown\Phi_{k,p}^{(m)}(M) \\ &= (\Phi_{k,p}^{(m)}(M))^{1-p} \binom{m}{k}^{-1} U \cdot \operatorname{diag}\left(\delta_{1,k}(\lambda(M)), \ldots, \delta_{m,k}(\lambda(M))\right) \cdot U^T \end{aligned} \tag{6}$$

where U is a matrix, for which $M = U \operatorname{diag}(\lambda(M))U^T$.

Proof. First we will calculate the gradient of $f = \varphi_p^{\binom{m}{k}} \circ L_k^{(m)}$.

$$\begin{aligned} &\frac{\partial f(\lambda(M))}{\partial \lambda_l(M)} \\ &= \frac{1}{p} \left(\binom{m}{k}^{-1} \sum_{1 \leq i_1 < \ldots < i_k \leq m} \left(\sum_{j=1}^{k} \lambda_{i_j}(M) \right)^{p} \right)^{\frac{1}{p}-1} p \binom{m}{k}^{-1} \delta_{l,k}(\lambda(M)) \\ &= \left(\varphi_p^{\binom{m}{k}} \circ L_k^{(m)}(\lambda(M)) \right)^{1-p} \binom{m}{k}^{-1} \delta_{l,k}(\lambda(M)). \end{aligned}$$

Therefore

$$\triangledown f(\lambda(M)) = \left(\varphi_p^{\binom{m}{k}} \circ L_k^{(m)}(\lambda(M))\right)^{1-p} \binom{m}{k}^{-1} (\delta_{1,k}(\lambda(M)), \ldots, \delta_{m,k}(\lambda(M)))^T .$$

Now we can calculate the gradient of Φ according to Lemmas 3.2 and 3.3:

$$\begin{aligned}\triangledown \Phi_{k,p}^{(m)}(M) &= U(\operatorname{diag} \triangledown f(\lambda(M)))U^T \\ &= (\Phi_{k,p}^{(m)}(M))^{1-p} \binom{m}{k}^{-1} U \operatorname{diag}(\delta_{1,k}(\lambda(M)), \ldots, \delta_{m,k}(\lambda(M)))\, U^T.\end{aligned}$$

□

3.2. Boundaries for $\Phi_{E_k}^{(m)}$-optimal values

In this subsection, we prove some relationships between $\Phi_{E_k}^{(m)}$ and $\Phi_{k,p}^{(m)}$ criteria, which will be used to numerically determine the boundaries for $\Phi_{E_k}^{(m)}$-efficiencies based on $\Phi_{k,p}^{(m)}$-efficiencies.

Theorem 3.3. *Let* $k \in \{1, \ldots, m\}$ *and let* $\operatorname{rank}(M) \geq m - k + 1$. *Then for all* $p \in (-\infty, 0)$*:*

$$1 \leq \frac{\Phi_{k,p}^{(m)}(M)}{\Phi_{E_k}^{(m)}(M)} \leq \binom{m}{k}^{-1/p}. \tag{7}$$

Proof. Let $n = \binom{m}{k}$ and $p < 0$. Let $a_j = \sum_{i \in I_j} \lambda_i(M)$ for $j = 1, \ldots, n$, where $I_1, \ldots, I_n$ are all subsets of $\{1, \ldots, m\}$ with k elements and $I_1 = \{1, \ldots, k\}$. Note that $a_1 \leq a_j$ for all j and $a_1 > 0$, because $\operatorname{rank}(M) \geq m - k + 1$.
Note also that

$$\begin{aligned}\frac{\Phi_{k,p}^{(m)}(M)}{\Phi_{E_k}^{(m)}(M)} &= \binom{m}{k}^{-1/p} \left(\frac{\displaystyle\sum_{1 \leq i_1 < \ldots < i_k \leq m} \left(\sum_{j=1}^{k} \lambda_{i_j}(M) \right)^p}{\left(\displaystyle\sum_{i=1}^{k} \lambda_i(M) \right)^p} \right)^{1/p} \\ &= \binom{m}{k}^{-1/p} \left(1 + \left(\frac{a_2}{a_1}\right)^p + \ldots + \left(\frac{a_n}{a_1}\right)^p \right)^{1/p}.\end{aligned} \tag{8}$$

At the same time we have

$$\binom{m}{k}^{1/p} \leq \left(1 + \left(\frac{a_2}{a_1}\right)^p + \ldots + \left(\frac{a_n}{a_1}\right)^p \right)^{1/p} \leq 1 \tag{9}$$

because $0 \leq \left(\frac{a_j}{a_1}\right)^p \leq 1$ for all $j = 1, \ldots, n$ and $p < 0$. Combining (8) and (9) we get both inequalities of the theorem. □

Note that the Theorem 3.3 implies that $\Phi^{(m)}_{k,p}$ converges for $p \to -\infty$ to $\Phi^{(m)}_{E_k}$ uniformly on the set of all information matrices.

More importantly, we have the following corollary:

Corollary 3.1. $\forall M \in \mathcal{S}^m_+, \ \ \forall k \in \{1, \dots, m\}, \ \ \forall p \in (-\infty, 0)$*:*

$$\mathrm{eff}(M \mid \Phi^{(m)}_{E_k}) \geq \binom{m}{k}^{1/p} \mathrm{eff}(M \mid \Phi^{(m)}_{k,p}).$$

Proof. If $\mathrm{eff}(M \mid \Phi^{(m)}_{k,p}) = 0$, the proof is trivial.
Let $\mathrm{eff}(M \mid \Phi^{(m)}_{k,p}) > 0$. Then $\Phi^{(m)}_{k,p}(M) > 0$, i.e. $\mathrm{rank}(M) \geq m - k + 1$. Let $M_{k,p}$ be a $\Phi^{(m)}_{k,p}$-optimal matrix and let M_{E_k} be a Φ_{E_k}-optimal matrix. Then:

$$\begin{aligned}\frac{\mathrm{eff}(M \mid \Phi^{(m)}_{E_k})}{\mathrm{eff}(M \mid \Phi^{(m)}_{k,p})} &= \frac{\Phi^{(m)}_{E_k}(M)}{\Phi^{(m)}_{E_k}(M_{E_k})} \cdot \frac{\Phi^{(m)}_{k,p}(M_{k,p})}{\Phi^{(m)}_{k,p}(M)} \\ &= \frac{\Phi^{(m)}_{E_k}(M)}{\Phi^{(m)}_{k,p}(M)} \cdot \frac{\Phi^{(m)}_{k,p}(M_{k,p})}{\Phi^{(m)}_{k,p}(M_{E_k})} \cdot \frac{\Phi^{(m)}_{k,p}(M_{E_k})}{\Phi^{(m)}_{E_k}(M_{E_k})} \geq \binom{m}{k}^{1/p}\end{aligned}$$

since the first factor in this product is greater than or equal to $\binom{m}{k}^{1/p}$ from the right inequality of Theorem 3.3, the second factor is greater than or equal to 1, because $M_{k,p}$ is $\Phi_{k,p}$-optimal, and the third factor is greater than or equal to 1 from the left inequality of Theorem 3.3. □

4. Algorithm for computing the boundaries of $\Phi^{(m)}_{E_k}$-optimal values

The following theorem tells us that for finding a design whose $\Phi^{(m)}_{E_k}$-efficiency is at least α, we only need to find a design whose $\Phi^{(m)}_{k,p}$-efficiency is at least $\beta > \alpha$, where $p \leq \ln_{\alpha/\beta} \binom{m}{k}$. But the criterion $\Phi^{(m)}_{k,p}$ is differentiable, therefore such design can be constructed using standard iterative algorithms.

Theorem 4.1. *If* $\mathrm{eff}(M(\xi) \mid \Phi^{(m)}_{k,p}) \geq \beta$ *for some design* ξ *and* $p \leq \ln_{\alpha/\beta} \binom{m}{k}$, $p \in (-\infty, 0)$, *then* $\mathrm{eff}(M(\xi) \mid \Phi^{(m)}_{E_k}) \geq \alpha$

Proof. We can easily see that $p \leq \ln_{\alpha/\beta} \binom{m}{k}$ implies $\alpha \leq \beta \binom{m}{k}^{1/p}$.
Using Corollary 3.1 we obtain

$$\mathrm{eff}(M(\xi) \mid \Phi^{(m)}_{E_k}) \geq \binom{m}{k}^{1/p} \mathrm{eff}(M(\xi) \mid \Phi^{(m)}_{k,p}) \geq \binom{m}{k}^{1/p} \beta \geq \alpha.$$

□

Now we can proceed to constructing the boundaries for $\Phi^{(m)}_{E_k}$-optimal values. In the first step, we will iteratively compute $\Phi^{(m)}_{k,p}$-optimal design.

The following algorithm is based on the algorithms described in [**8**, chapter V]. As an input, we need:

- the required $\Phi_{E_k}^{(m)}$ efficiency α, from which using the Theorem 4.1, we get β and p
- model function f
- experimental domain $\mathcal{X}$
- starting information matrix $M_1 \in \mathcal{S}_{++}^m$
- k (the parameter of $\Phi_{E_k}^{(m)}$)

The sequence of the information matrices $\{M_i\}$ is constructed in the following way:

1. let $i = 1$
2. if
$$\frac{\Phi_{k,p}^{(m)}(M_i)}{\max_{x\in\mathcal{X}} f^T(x) \bigtriangledown \Phi_{k,p}^{(m)}(M_i) f(x)} > \beta,$$
 stop
 else go to 4.
3. compute the information matrix M_{i+1} according to the formula
$$M_{i+1} = \frac{i}{i+1} M_i + \frac{1}{i+1} f(x_{i+1}) f^T(x_{i+1})$$
 where $x_{i+1} = \arg\max_{x\in\mathcal{X}} f^T(x) \bigtriangledown \Phi_{k,p}^{(m)}(M_i) f(x)$
 let $i = i + 1$
4. go to 3.

Note: If the inequality in 2. holds, then $\mathrm{eff}(M(\xi_i) \mid \Phi_{k,p}^{(m)}) > \beta$.

Let M be a matrix obtained by iterative computation of the $\Phi_{k,p}^{(m)}$-optimal design described above. Then the lower and upper boundaries for $\Phi_{E_k}^{(m)}$-optimal values are as follows:

- lower boundary: $\Phi_{E_k}^{(m)}(M)$
- 1st upper boundary: $\Phi_{k,p}^{(m)}(M)/\beta$ follows from

$$\begin{aligned} \mathrm{eff}(M \mid \Phi_{k,p}^{(m)}) \geq \beta \quad &\Rightarrow \quad \frac{\Phi_{k,p}^{(m)}(M)}{\Phi_{k,p}^{(m)}(M_{k,p})} \geq \beta \\ &\Rightarrow \quad \Phi_{k,p}^{(m)}(M_{k,p}) \leq \frac{\Phi_{k,p}^{(m)}(M)}{\beta}, \end{aligned} \tag{10}$$

 where $M_{k,p}$ is a $\Phi_{k,p}^{(m)}$-optimal matrix. Therefore:

$$\Phi_{E_k}^{(m)}(M) \leq \Phi_{E_k}(M_{E_k}^{(m)}) \leq \Phi_{k,p}^{(m)}(M_{E_k}) \leq \Phi_{k,p}^{(m)}(M_{k,p}) \leq \frac{\Phi_{k,p}^{(m)}(M)}{\beta}$$

- 2nd upper boundary:

$$\max f^T(x) Y f(x),$$

where Y is the subgradient of $\Phi_{E_k}^{(m)}$ in the point M obtained as follows: $Y = \sum_{i=1}^{k} u_i u_i^T$, u_i is an eigenvector of M corresponding to the eigenvalue $\lambda_i(M)$ (it follows from the form of the subdifferential of Φ_{E_k} and the theorem on the boundary of $\Phi_{E_k}^{(m)}$-efficiency in [**5**])

5. Example: Polynomial model on $\langle -1, 1\rangle$

The polynomial model for the measurement in the point x is defined as

$$y(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \ldots + \theta_d x^d + \varepsilon$$

where d is the degree of the model and $\theta_0, \ldots, \theta_d$ are the parameters of the model, i.e. the number of parameters is $m = d+1$. Thus we have $f(x) = (1, x, x^2, \ldots, x^d)$. Suppose the measurements are carried out on the set $\mathcal{X} = \langle -1, 1\rangle$. We will estimate the $\mathbb{O}$-minimal efficiency of the D-optimal, E-optimal and uniform arcsine designs.

Note: The uniform arcsine design is understood to be the uniform design on the points of the arcsine support in the sense of the definition in [**9**, p. 217]. The construction of D-optimal design is based on [**9**, Chapter 9]. The construction of E-optimal design follows from the theorems in [**10**] and [**9**, Part 9.13].

The results were obtained in the following way:

Table 1. Lower and upper boundaries for $\Phi_{E_k}^{(m)}$-optimal values for the degrees $d = 2, \ldots, 8$ of polynomial regression and $k = 1, \ldots, d+1$. Optimum values are indicated by bold face. The upper boundaries are computed as the minimum of the 1^{st} and 2^{nd} upper boundaries in (10).

k/degree	2		3		4		5	
1	0.194595	0.201389	0.039198	0.042363	0.007482	0.008072	0.001413	0.001577
2	0.976261	1.043871	0.192313	0.201618	0.039069	0.042727	0.007562	0.008249
3	2.852	**3**	1.98444	2.208806	0.32073	0.347491	0.082466	0.090713
4			3.804006	**4**	1.957065	2.110195	0.31006	0.354067
5					4.754523	**5**	2.978997	3.346387
6							5.704555	**6**

k/degree	6		7		8	
1	0.000263	0.000287	4.86E-05	5.43E-05	8.83E-06	9.86E-06
2	0.001426	0.001579	0.000267	0.000283	4.75E-05	5.31E-05
3	0.018261	0.019762	0.003891	0.004303	0.000807	0.000891
4	0.083396	0.090054	0.018156	0.020106	0.00389	0.00431
5	0.408581	0.45114	0.127203	0.1391	0.03028	0.032078
6	2.940145	3.163073	0.428754	0.468567	0.126843	0.139165
7	6.654964	**7**	3.977078	4.469847	0.509199	0.53346
8			7.607143	**8**	3.918662	4.218915
9					8.565743	**9**

Table 2. Lower and upper boundaries for the minimal efficiencies of arcsine, D-optimal, and E-optimal designs for various degrees of polynomial regression.

k/degree	2		3		4		5	
arcsine	0.7257	0.751	0.5573	0.6202	0.5858	0.6257	0.4683	0.526
D-optim	0.7257	0.751	0.5513	0.6137	0.5723	0.6078	0.4505	0.506
E-optim	0.5747	0.6145	0.3589	0.3995	0.4104	0.4425	0.3142	0.3529

k/degree	6		7		8	
arcsine	0.502	0.528	0.4168	0.4685	0.4408	0.4632
D-optimal	0.4801	0.5049	0.3966	0.4457	0.4189	0.4401
E-optimal	0.3525	0.3792	0.2838	0.319	0.3095	0.3276

- we computed the D-optimal design ξ_D^*, E-optimal design ξ_E^* and uniform arcsine design ξ_A^*
- using the eigenvalues of $M(\xi_D^*)$, $M(\xi_E^*)$ and $M(\xi_A^*)$ we determined the values of $\Phi_{E_k}^{(m)}$ criteria
- then we calculated the $\Phi_{E_k}^{(m)}$-optimal design with the help of the Theorem 4.1 (putting $\beta = 0.95$ and $\alpha = 0.9$) and determined boundaries for $\Phi_{E_k}^{(m)}$-optimal values (Table 1)
- finally, we constructed the intervals containing the minimal efficiencies ξ_D^*, ξ_E^*, ξ_A^* according to $\mathbb{O}$ (Table 2)

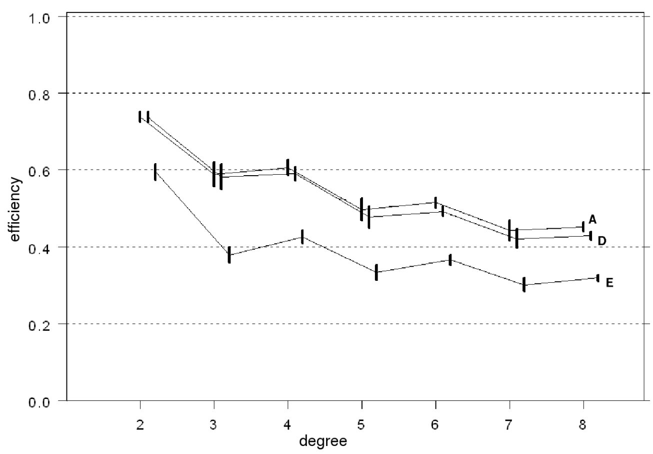

Figure 1. The boundaries for minimal efficiencies of the arcsine, D-optimal and E-optimal designs for various degrees of polynomial regression. The vertical lines denote the interval from the lower to the upper boundary of the corresponding minimal efficiency. To improve readability, the graphs of individual design are slightly shifted.

Note: The minimal efficiencies for the degrees 2, 3, 4, which were computed precisely in [**3**], are in agreement with our numerical boundaries in the Table 2.

References

1. Bondar J. V., *Universal optimality of experimental designs: definitions and a criterion*, The Canadian Journal of Statistics **11(4)** (1983), 325–331.

2. Borwein J. M., Lewis A. S., *Convex analysis and nonlinear optimization, Theory and examples*, CMS Books in Mathematics, Springer-Verlag, New York 2000.

3. Harman R., *Minimal efficiency of designs under the class of orthogonally invariant information criteria*, Metrika **60** (2004), 137–153.

4. ———, *Lower bounds on efficiency ratios based on Φ_p-optimal designs,* Proceedings from the conference MODA7, Kapellerput Heeze 2004.

5. ———, *Minimálna eficiencia experimentálnych návrhov v triede ortogonálne invariantných kritérií*, Dizertačná práca, FMFI UK 2004.

6. Horn R. A. and Johnson C. R., *Matrix analysis*, Cambridge University Press, Cambridge 1985.

7. Lewis A. S., C*onvex analysis on the Hermitian matrices*, SIAM Journal of Optimization **6** (1996), 164–177.

8. Pázman A., *Foundations of optimum experimental design*, D. Reidel, Dordrecht 1986.

9. Pukelsheim F. *Optimal design of experiments*, John Wiley & Sons, New York 1993.

10. Pukelsheim F., Studden W. J., *E-optimal designs for polynomial regression,* Annals of Statistics **21(1)** (1993), 402–415.

11. Rockafellar R. T., *Convex analysis*, Princeton University Press, Princeton 1970.

L. Bušová, Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava,
e-mail: `Lenka.Busova@fmph.uniba.sk`