Applications of necessary and sufficient conditions for maximin efficient designs

Christine H. Müller, Georg-August-University, Göttingen
and
Andrej Pázman, Comenius University, Bratislava [1]

March 17, 1998

Abstract

General sufficient and necessary conditions for minimax design are here reconsidered in a form allowing application in various optimal design problems. In combination with the Elfving theorem they are used to find maximin efficient designs for a two-dimensional linear extrapolation, and to find the optimum design for estimating the maximum point of a quadratic response function with intercept. An alternative proof of a recently published relation between D-optimality and maximin efficiency is given. It is shown that for exponential growth curve models with one parameter, maximin efficient designs can not be one point designs. A similar result is obtained for growth curve models with two parameters.

Keywords: Optimal design, equivalence theorem, nonlinear models and aspects, position of extremes, exponential growth model.

AMS subject classification: 62K05

1 Introduction

We consider a general linear or nonlinear regression model

$$Y_n = \eta(t_n, \theta) + Z_n,\ n = 1, \ldots, N,$$

with unknown $\theta \in \Theta \subset I\!R^p$, where the observations $Y_1, \ldots, Y_N$ are made at design points $t_1, \ldots, t_N \in \mathcal{T}$. For finding an optimal design for estimating θ or an aspect $\varphi(\theta)$ it is usually useful to extend the exact design $d = (t_1, \ldots, t_N)$ to a design measure $\delta_N = \frac{1}{N}\sum_{n=1}^{N} e_{t_n}$, where e_t denotes the one-point (Dirac) measure on t. Then

$$\Delta = \{\delta;\ \delta \text{ is probability measure on } \mathcal{T}\}$$

[1]Research supported by the grant 436 SLK 17/8/96 of the Deutsche Forschungsgemeinschaft and by the Slovak VEGA grant 1/4196/97

is the set of all possible designs.

Often designs are judged by an convex criterion, which under circumstances depends on some parameter $\gamma \in \Gamma$. So, here we assume that $\Psi_\gamma : \Delta \to I\!R$ is for all $\gamma \in \Gamma$ a convex design criterion and that an optimal design should minimize the maximum with respect to γ, that is, an optimal design is a minimax design in the following sense.

Definition 1.1 *$\delta^* \in \Delta$ is a minimax design with respect to Ψ_γ in Δ if and only if*

$$\delta^* \in \arg\min\{\max_{\gamma\in\Gamma} \Psi_\gamma(\delta);\ \delta \in \Delta\}.$$

For example, E-optimal designs in a linear model given by $\eta(t,\theta) = f(t)^\top\theta$ with $f : T \to I\!R^p$ and $\theta \in I\!R^p$ are minimax solutions in the above sense. Recall that a E-optimal design minimizes the maximum eigenvalue of the inverse of the information matrix $\mathcal{I}(\delta) = \int f(t)\, f(t)^\top \delta(dt)$. Hence, setting $\Gamma = \{\gamma \in I\!R^p;\ \gamma^\top\gamma = 1\}$ and $\Psi_\gamma(\delta) = \gamma^\top \mathcal{I}(\delta)^{-1}\gamma$, E-optimal designs are minimax designs.

But also maximin efficient designs are minimax solutions in the above sense. Namely, setting $\Gamma = \Theta$ and $\Psi_\theta(\delta) = \frac{\Phi_\theta(\delta)}{\Phi_\theta(\delta_\theta)}$ for $\gamma = \theta \in \Theta$, where $\Phi_\theta(\delta)$ is a a given optimality criterion and

$$\delta_\theta \in \arg\min\{\Phi_\theta(\delta);\ \delta \in \Delta\}$$

is the locally optimal design. A maximin efficient design δ^* satisfies

$$\delta^* \in \arg\max\{\min_{\theta\in\Theta} \frac{\Phi_\theta(\delta_\theta)}{\Phi_\theta(\delta)};\ \delta \in \Delta\},$$

so it is a minimax design with this choice of Γ and of $\Psi_\theta(\delta)$. Maximin efficient designs are particularly important for estimating nonlinear aspects $\varphi(\theta)$ in linear models and for estimating θ or aspects of θ in nonlinear models. See e.g. Silvey (1980), Ford et al. (1989) and McDonald and Studden (1990). In the case of a linear model given by $\eta(t,\theta) = f(t)^\top\theta$, the criterion Φ_θ is a functional of $\dot\varphi_\theta^\top \mathcal{I}(\delta)^{-1}\dot\varphi_\theta$, where $\dot\varphi_\theta = \frac{\partial}{\partial\theta}\varphi(\theta)$ and $\mathcal{I}(\delta) = \int f(t)f(t)^\top\delta(dt)$. In the case of a nonlinear model, Φ_θ is a functional of $\mathcal{I}_\theta(\delta)^{-1}$, where $\mathcal{I}_\theta(\delta) = \int f_\theta(t) f_\theta(t)^\top \delta(dt)$ and $f_\theta(t) = \frac{\partial}{\partial\theta}\eta(t,\theta)$.

In Section 2 we provide general necessary and sufficient conditions for minimax design. These conditions are applied in Section 3 to maximin efficient designs in linear models and in Section 4 to maximin efficient designs in nonlinear model. In Section 3 in particular an alternative proof of a result of Schwabe (1997) concerning the relationship between maximin efficient designs and D-optimal designs is given. Moreover, using the sufficient condition, maximin efficient designs are derived for situations where the assumptions of Schwabe (1997) are not satisfied, namely for two-dimensional linear extrapolation and for estimating the maximum point of a quadratic response function with intercept. In Section 4 the necessary condition is used to show that in a one-parameter exponential growth curve model one-point designs are not maximin efficient. A similar result is also shown for the two-parameter growth curve model.

2 Necessary and sufficient conditions for minimax designs

A necessary and sufficient condition, related to the one which will be applied here, has been presented in Fedorov (1980), however, with just an indication of the proof. We present a modification with declared assumptions and a detailed direct proof.

We shall say that a design δ^* follows the assumption (A) if for any design $\delta \in \Delta$ there is a $c > 0$ such that the functions

$$(\epsilon, \gamma) \in [0, c] \times \Gamma \to \Psi_\gamma\left((1-\epsilon)\delta^* + \epsilon\delta\right)$$

and

$$(\epsilon, \gamma) \in [0, c] \times \Gamma \to \frac{d\Psi_\gamma\left((1-t)\delta^* + t\delta\right)}{dt} \Big|_{t=\epsilon}$$

are continuous.

Denote by $\nabla_\gamma(\delta^*, \delta)$ the directional derivative of $\Psi_\gamma(\delta)$ in the point δ^* and in the direction δ

$$(2.1) \quad \nabla_\gamma(\delta^*, \delta) := \lim_{\epsilon \downarrow 0} \frac{\Psi_\gamma((1-\epsilon)\delta^* + \epsilon\delta) - \Psi_\gamma(\delta^*)}{\epsilon}.$$

Theorem 2.1 *Let δ^* be such that the set $\mathcal{M} = \arg\max_{\gamma\in\Gamma} \Psi_\gamma(\delta^*)$ is finite.*
(i) If there exists $\lambda : \mathcal{M} \to [0, 1]$ with $\sum_{\gamma\in\mathcal{M}} \lambda(\gamma) = 1$ and

$$\sum_{\gamma\in\mathcal{M}} \lambda(\gamma)\nabla_\gamma(\delta^*, \delta) \geq 0$$

for all $\delta \in \Delta$, then the design δ^* is a minimax design with respect to Ψ_γ in Δ.
(ii) The converse implication is true, if assumption (A) holds and if the set Γ is compact.

Proof. Denote by

$$Q(\delta) = \max_{\gamma\in\Gamma} \Psi_\gamma(\delta)$$

the minimax optimality criterion. Note that $Q(\delta)$ is a convex function, since $\Psi_\gamma(\delta)$ is convex for every γ. Hence, its directional derivative,

$$\nabla Q(\delta^*, \delta) = lim_{n\to\infty} \frac{Q(\delta_n) - Q(\delta^*)}{1/n}$$

exists. Here $\delta_n = (1 - (1/n))\delta^* + (1/n)\delta$.

1) Proof of sufficiency: Take any δ. Take any $\gamma^* \in \mathcal{M}$. By the definition of $Q(\delta)$ and of γ^* we have

$$\frac{Q(\delta_n) - Q(\delta^*)}{(1/n)} \geq \frac{\Psi_{\gamma^*}(\delta_n) - \Psi_{\gamma^*}(\delta^*)}{(1/n)}$$

Taking $n \to \infty$ on both sides, we obtain

$$\nabla Q(\delta^*, \delta) \geq \nabla_{\gamma^*}(\delta^*, \delta)$$

for every $\gamma^* \in \mathcal{M}$. Hence for any δ we have

$$\nabla Q(\delta^*, \delta) \geq \sum_{\gamma \in \mathcal{M}} \lambda(\gamma) \nabla_\gamma(\delta^*, \delta) \geq 0$$

which implies that δ^* is Q-optimal, hence minimax.

2) Proof of necessity: Let

$$\mathcal{M} = \{\gamma_1, ..., \gamma_k\}.$$

Fix any δ and define δ_n as above. Choose some $\gamma(n) \in \arg\max_{\gamma \in \Gamma} \Psi_\gamma(\delta_n)$. Since Γ is compact, the sequence $\gamma(n)$, $n = 1, 2, ...$, has a convergent subsequence, and with no lack of generality, we can suppose that the whole sequence is converging to a point $\tilde{\gamma} \in \Gamma$. We can write

$$0 \leq \nabla Q(\delta^*, \delta) \leq \frac{Q(\delta_n) - Q(\delta^*)}{1/n} \leq \frac{\Psi_{\gamma(n)}(\delta_n) - \Psi_{\gamma(n)}(\delta^*)}{1/n}$$

$$= \frac{d\Psi_{\gamma(n)}\left((1-t)\delta^* + t\delta\right)}{dt} \,\Big|_{t=t_n}$$

for some $t_n \in [0, 1/n]$. The first inequality holds, because δ^* is Q-optimal. The second follows from the definition of the directional derivative and the convexity of $Q(\delta)$ (cf. e.g. Pázman 1986, page 113). The third follows from the definition of $Q(.)$ and of $\gamma(n)$. The last equality follows from the mean value theorem.

Now take the limit of the last term for $n \to \infty$. ¿From the second continuity assumption in (A) we obtain

$$0 \leq \nabla_{\tilde{\gamma}}(\delta^*, \delta).$$

On the other hand

$$\tilde{\gamma} = \gamma_i \tag{2.2}$$

for some i. This can be seen from the following relations which hold for some $\gamma_i \in \mathcal{M}$:

$$Q(\delta^*) = \Psi_{\gamma_i}(\delta^*) = \lim_{n \to \infty} \Psi_{\gamma_i}(\delta_n) \leq \lim_{n \to \infty} \Psi_{\gamma(n)}(\delta_n) \overset{(1)}{=} \Psi_{\tilde{\gamma}}(\delta^*) \leq Q(\delta^*).$$

In particular, the equality (1) follows from the first continuity assumption in (A).

Now define

$$H = \{(x_1, \ldots, x_k) \in I\!R^k;\ \text{there exists } \delta \in \Delta \text{ with } x_i \geq \nabla_{\gamma_i}(\delta^*, \delta) \text{ for } i = 1, ..., k\}.$$

¿From the convexity of $\Psi_\gamma(\delta)$ and from (2.1) we have that $\nabla_{\gamma_i}(\delta^*, \delta)$ is convex in δ. Hence, the set H is also convex. Further, according to (2.2) we have that for every δ at least for one i

$$\nabla_{\gamma_i}(\delta^*, \delta) \geq 0.$$

Hence the set H has no common points with the interior of the convex set

$$Z = \{(z_1, \ldots, z_k) \in I\!R^k;\ z_i \leq 0 \text{ for } i = 1, ..., k\}.$$

However, the zero vector is in the intersection of Z and H. Applying the separating hyperplane theorem, we obtain that there are numbers $\mu_1, ..., \mu_k$ such that

$$\sum_{i=1}^{k} \mu_i \nabla_{\gamma_i}(\delta^*, \delta) \geq 0 \text{ for every } \delta \in \Delta \tag{2.3}$$

and

$$\sum_{i=1}^{k} \mu_i z_i \leq 0 \text{ for every } (z_1, \ldots, z_k) \in Z.$$

By taking negative unit vectors $-u_i \in Z$ we obtain that $\mu_i \geq 0$. Taking finally $\lambda_i = \frac{\mu_i}{\sum_j \mu_j}$ we obtain from (2.3) the required converse statement. □

The sufficient condition can be generalized to the case that $\mathcal{M} = \arg\max_{\gamma\in\Gamma} \Psi_\gamma(\delta^*)$ is infinite. In that case it is sufficient for maximin efficiency of a design δ^* that a probability measure λ on $\mathcal{M}$ exists with

$$\int \nabla_\gamma(\delta^*, \delta)\lambda(d\gamma) \geq 0$$

for all $\delta \in \Delta$. The proof is the same as for finite $\mathcal{M}$. The necessary condition can be given for infinite $\mathcal{M}$ as well. However, in applications the set $\mathcal{M}$ is finite (see the examples below). Note that the sufficient condition can be also shown by the wellknown relationship between minimax solutions and Bayesian solutions with respect to least favorable priors (see e.g. Schervish 1995, Theorem 3.65, page 168).

In the particular case that $\Gamma = C \subset I\!R^p$, $\Psi_c(\delta) = a(c)\ c^\top \mathcal{I}(\delta)^{-1}c$ for $c = \gamma \in C$ with $a : C \to I\!R$, Theorem 2.1 can be given more explicitely.

Theorem 2.2 *Let $\mathcal{I}(\delta^*)$ be regular and $\mathcal{M} = \arg\max_{c\in C} a(c)\ c^\top \mathcal{I}(\delta^*)^{-1}c$ be a finite set. If $\lambda : \mathcal{M} \to [0, 1]$ with $\sum_{c\in\mathcal{M}} \lambda(c) = 1$ and*

$$\sum_{c\in\mathcal{M}} \lambda(c) \frac{(c^\top \mathcal{I}(\delta^*)^{-1} f(t))^2}{c^\top \mathcal{I}(\delta^*)^{-1} c} \leq 1 \tag{2.4}$$

for all $t \in T$, then δ^* is a minimax design in Δ. The converse is true if C is compact and $a(c)$ is continuous.

Proof. The directional derivative of $\Phi_c(\delta) = c^\top \mathcal{I}(\delta)^{-1} c$ is known (see e.g. Silvey 1980, page 49). In particular this derivative is linear in δ, so that only derivatives in directions of one-point (Dirac) measures e_t with $t \in T$ have to be regarded. Since for fixed c the multiplicator $a(c)$ is only a constant, the derivative of Ψ_c at δ^* is

$$\nabla_c(\delta^*, e_t) = a(c) \ \left(c^\top \mathcal{I}(\delta^*)^{-1} c - \left(c^\top \mathcal{I}(\delta^*)^{-1} f(t) \right)^2 \right).$$

Because for all $c \in \mathcal{M}$ the value of $\Psi_c(\delta^*) = a(c) \ c^\top \mathcal{I}(\delta^*)^{-1} c$ is constant, Theorem 2.1 provides the proof of sufficiency. The continuity assumptions (A) follow from the expressions for $\Psi_c(\delta)$ and $\nabla_c(\delta^*, \delta)$ so that the necessity of (2.4) holds as well. □

Example 2.1. (E-optimal designs for linear regression) Regard the model $Y = f(t)^\top \theta + Z$, where $\theta \in I\!R^2$ and $f(t) = (1, t)^\top \in I\!R^2$ with $t \in \mathcal{T} = [-b, b]$ for $b > 0$. Theorem 2.2 provides now an alternative proof of the E-optimality of the design $\delta^* = \frac{1}{2} e_{-b} + \frac{1}{2} e_b$. Setting $c = (c_1, c_2)^\top$, where $c_1^2 = 1 - c_2^2$, we have

$$c^\top \mathcal{I}(\delta^*)^{-1} c = c_1^2 + \frac{c_2^2}{b^2} = 1 + c_2^2 \left(\frac{1}{b^2} - 1 \right).$$

Hence for $b^2 \geq 1$

$$(1, 0)^\top \in \arg \max_{c:|c|=1} c^\top \mathcal{I}(\delta^*)^{-1} c$$

and

$$(1, 0) \ \mathcal{I}(\delta^*)^{-1} f(t) = 1 = \max_{c:|c|=1} c^\top \mathcal{I}(\delta^*)^{-1} c$$

for all $t \in [-b, b]$. Similarly, for $b^2 < 1$

$$(0, 1)^\top \in \arg \max_{c:|c|=1} c^\top \mathcal{I}(\delta^*)^{-1} c$$

and

$$|(0, 1) \ \mathcal{I}(\delta^*)^{-1} f(t)|^2 = \left| \frac{t}{b^2} \right|^2 \leq \frac{b^2}{b^4} = \max_{c:|c|=1} c^\top \mathcal{I}(\delta^*)^{-1} c$$

for all $t \in [-b, b]$, so that the conditions of Theorem 2.2 are satisfied. □

3 Maximin efficiency in linear models

In this section we assume $\Gamma = C \subset I\!R^p$, $\Psi_c(\delta) = \frac{c^\top \mathcal{I}(\delta)^{-1} c}{c^\top \mathcal{I}(\delta_c)^{-1} c}$ for $c \in C$, where

$$\delta_c \in \arg \min \{ c^\top \mathcal{I}(\delta)^{-1} c; \ \delta \in \Delta \}$$

is the locally optimal design. For checking the conditions of Theorem 2.2 we must not regard the whole set C for determining the maximum points of $\Psi_c(\delta^*)$. It is sufficient to regard a set G. For defining the set G let be $\mathcal{E} := \{f(t);\ t \in T\} \cup \{-f(t);\ t \in T\}$ the Elfving set, and $\overline{C}$ the closure of the set $\{\lambda c \in I\!R;\ \lambda \in I\!R,\ c \in C\}$. Then the set G is the intersection of $\overline{C}$ and the boundary of $\mathcal{E}$. For this set the theorem of Elfving (1952) provides $c^\top \mathcal{I}(\delta_c)^{-1} c = 1$ for all $c \in G$ and in particular

$$(3.5) \qquad \max_{c \in C} \Psi_c(\delta) = \max_{c \in \overline{C}} \Psi_c(\delta) = \max_{c \in G} c^\top \mathcal{I}(\delta)^{-1} c$$

for all δ.

With this property Theorem 2.2 provides an alternative proof of Theorem 2.2 (i) in Schwabe (1996), which shows that under a simple condition D-optimal designs are maximin efficient.

Theorem 3.1 *If δ^* is D-optimal in Δ and has a finite support supp(δ^*) and $\{f(t);\ t \in$ supp$(\delta^*)\} \subset \overline{C}$, then δ^* is maximin efficient in Δ.*

Proof. Let be δ^* D-optimal. The minimum ellipsoid argument of Silvey (1972) and Sibson (1972) (see Silvey 1980, page 41) provides that the ellipsoid $\{c \in I\!R^p;\ c^\top \mathcal{I}(\delta^*)^{-1} c \leq p\}$ is the smallest ellipsoid containing the Elfving set $\mathcal{E}$. Moreover, the intersection of this ellipsoid with the boundary of $\mathcal{E}$ is $\{f(\tau);\ \tau \in \mathrm{supp}(\delta^*)\} \cup \{-f(\tau);\ \tau \in \mathrm{supp}(\delta^*)\}$. This means that $\mathcal{M} := \{f(\tau);\ \tau \in \mathrm{supp}(\delta^*)\} \subset \arg\max_{c \in \overline{C}} \Psi_c(\delta^*)$ and $c^\top \mathcal{I}(\delta^*)^{-1} c = p$ for all $c \in \mathcal{M}$. For $c = f(\tau) \in \mathcal{M}$ set $\lambda(f(\tau)) = \delta^*(\{\tau\})$. Then the equivalence theorem for D-optimality provides

$$\begin{aligned}
&\sum_{c \in \mathcal{M}} \lambda(c) \frac{(c^\top \mathcal{I}(\delta^*)^{-1} f(t))^2}{c^\top \mathcal{I}(\delta^*)^{-1} c} \\
&= \frac{1}{p} \sum_{\tau \in \mathrm{supp}(\delta^*)} \delta^*(\{\tau\}) f(t)^\top \mathcal{I}(\delta^*)^{-1} f(\tau)\, f(\tau)^\top \mathcal{I}(\delta^*)^{-1} f(t) \\
&= \frac{1}{p} f(t)^\top \mathcal{I}(\delta^*)^{-1} f(t) \\
&\leq 1
\end{aligned}$$

for all $t \in T$ so that according to Theorem 2.2 δ^* is maximin efficient. □

Several examples, where the conditions of Theorem 3.1 are satisfied, are given in Schwabe (1996). But using Theorem 2.2 and property (3.5) we can now derive maximin efficient designs in situations where the conditions of Theorem 3.1 are not satisfied.

3.1 Linear extrapolation

Assume a linear model $Y = f(t)^\top \theta + Z$, where $\theta \in I\!R^2$, and $f(t) = (1, t)^\top \in I\!R^2$ with $t \in \mathcal{T} = [0, 1]$. Of interest are all aspects $\varphi(\theta) = f(\tau)^\top \theta$ with $\tau > 1$. For fixed τ the best extrapolations designs

$$\delta_\tau := \arg\min\{f(\tau)^\top \mathcal{I}(\delta)^- f(\tau);\ \delta \in \Delta\}$$

are known from the theorem of Elfving (1952): Because

$$f(\tau) = \begin{pmatrix} 1 \\ \tau \end{pmatrix} = (\tau - 1)\begin{pmatrix} -1 \\ 0 \end{pmatrix} + \tau \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

the locally optimal design is $\delta_\tau = \frac{\tau-1}{2\tau-1}e_0 + \frac{\tau}{2\tau-1}e_1$.

Proposition 3.1 $\delta^* = \frac{1}{4}e_0 + \frac{3}{4}e_1$ *is the maximin efficient design, i.e.*

$$\delta^* = \arg\max\left\{\min_{\tau>1} \frac{f(\tau)^\top \mathcal{I}(\delta_\tau)^{-1} f(\tau)}{f(\tau)^\top \mathcal{I}(\delta)^{-1} f(\tau)};\ \delta \in \Delta\right\}.$$

Proof. Set $\mu := \frac{1}{2\tau-1}$. Then $\frac{1}{2\tau-1}(1,\tau)^\top = (\mu, \frac{\mu+1}{2})^\top$ and $G = \{(\mu, \frac{\mu+1}{2})^\top;\ \mu \in [0,1]\} \cup \{-(\mu, \frac{\mu+1}{2})^\top;\ \mu \in [0,1]\}$. Because

$$h(\mu) := \left(\mu, \frac{\mu+1}{2}\right) \mathcal{I}(\delta^*)^{-1} \left(\mu, \frac{\mu+1}{2}\right)^\top = \frac{4}{3}(\mu^2 - \mu + 1)$$

the maximum value $\frac{4}{3}$ of $h : [0,1] \to I\!R$ is attained by $\mu = 0$ and $\mu = 1$, i.e.

$$\left\{\begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\} \subset \arg\max_{c \in G} c^\top \mathcal{I}(\delta^*)^{-1} c.$$

Set $\lambda((0,\frac{1}{2})^\top) = \frac{1}{3}$ and $\lambda((1,1)^\top) = \frac{2}{3}$. Then

$$\begin{aligned}
&\lambda\left(\left(0,\frac{1}{2}\right)^\top\right)\left(\left(0,\frac{1}{2}\right)\mathcal{I}(\delta^*)^{-1}f(t)\right)^2 + \lambda\left((1,1)^\top\right)\left((1,1)\mathcal{I}(\delta^*)^{-1}f(t)\right)^2 \\
&= \frac{1}{3}\,\frac{4}{9}(4t-3)^2 + \frac{2}{3}\,\frac{16}{9}t^2 \\
&= \frac{4}{3\cdot 9}(24t^2 - 24t + 9) \\
&\le \frac{4}{3} = \max_{c\in G} c^\top \mathcal{I}(\delta^*)^{-1} c
\end{aligned}$$

for all $t \in [0,1]$ so that δ^* is maximin efficient according to Theorem 2.2. □

3.2 Maximum point of a quadratic function with intercept

Consider the quadratic regression model $Y = f(t)^\top\theta + Z$, where $f : [-1,1] \ni t \to (1,t,t^2)^\top \in I\!R^3$ and $\theta = (\theta_0,\theta_1,\theta_3)^\top \in \Theta := \{(\overline{\theta}_0,\overline{\theta}_1,\overline{\theta}_2)^\top \in I\!R^3;\ \theta_2 < 0\}$. The maximum point of the quadratic function $\theta_0 + \theta_1 t + \theta_2 t^2$ is $\varphi(\theta) = \frac{-\theta_1}{2\,\theta_2}$. For estimating the maximum point we can use $\varphi(\widehat{\theta})$, where $\widehat{\theta}$ is the least squares estimator for θ. Then $\varphi(\widehat{\theta})$ is asymptotically normally distributed with asymptotic covariance matrix $\dot{\varphi}_\theta^\top \mathcal{I}(\delta)^{-1}\dot{\varphi}_\theta$, where

$$\dot{\varphi}_\theta := \frac{\partial}{\partial\theta}\varphi(\theta) = \left(0, \frac{-1}{2\theta_2}, \frac{\theta_1}{2\,\theta_2^2}\right)^\top = \frac{-1}{2\,\theta_2}\left(0, 1, \frac{-\theta_1}{\theta_2}\right)^\top.$$

(See e.g. Silvey 1980, page 57). Hence $C = \{\dot{\varphi}_\theta;\ \theta \in \Theta\}$. If the quadratic function has no intercept θ_0, then the conditions of Theorem 3.1 are satisfied (see Schwabe 1996). But this is not the case if the quadratic function has an intercept. In Müller (1995) it was shown that $\frac{1}{3}(e_{-1} + e_0 + e_1)$ is maximin efficient in $\Delta^* = \{\delta;\ \mathrm{supp}(\delta) = \{-1, 0, 1\}\}$. We are now going to derive the maximin efficient design in $\Delta = \{\delta;\ \mathrm{supp}(\delta) \subset [-1, 1]\}$.

At first we construct the set G, the intersection of $\overline{C}$ and the boundary $\partial\mathcal{E}$ of the Elfving set $\mathcal{E}$. The set C is the union of the four sets $\{\dot{\varphi}_\theta;\ \frac{\theta_1}{\theta_2} \geq 1\}$, $\{\dot{\varphi}_\theta;\ 0 \leq \frac{\theta_1}{\theta_2} \leq 1\}$, $\{\dot{\varphi}_\theta;\ -1 \leq \frac{\theta_1}{\theta_2} \leq 0\}$, and $\{\dot{\varphi}_\theta;\ \frac{\theta_1}{\theta_2} \leq -1\}$.

For $\frac{\theta_1}{\theta_2} \geq 1$ we have with $\rho = \frac{\theta_2}{\theta_1} \in (0, 1]$

$$\begin{aligned}
\dot{\varphi}_\theta &= \frac{-\theta_1}{\theta_2^2}\left(0, \frac{\theta_2}{2\theta_1}, \frac{-1}{2}\right)^\top = \frac{-\theta_1}{\theta_2^2}\left(0, \frac{\rho}{2}, \frac{-1}{2}\right)^\top \\
&= \frac{-\theta_1}{\theta_2^2}\left(\frac{1}{2}(1,0,0)^\top + \frac{1}{2}\frac{1-\rho}{2}(-1,-1,-1)^\top + \frac{1}{2}\frac{1+\rho}{2}(-1,1,-1)^\top\right) \\
&= \frac{-\theta_1}{\theta_2^2}\left(\frac{1}{2}f(0)^\top - \frac{1}{2}\frac{1-\rho}{2}f(1)^\top - \frac{1}{2}\frac{1+\rho}{2}f(-1)^\top\right).
\end{aligned}$$

Hence, the intersection of $\partial\mathcal{E}$ and the closure of $\{\mu\,\dot{\varphi}_\theta;\ \mu > 0,\ \frac{\theta_1}{\theta_2} \geq 1\}$ is $G_1 := \{(0, \frac{\rho}{2}, \frac{-1}{2})^\top;\ \rho \in [0, 1]\}$. Note that a picture of this Elfving set is in Pukelsheim (1993), page 43.

For $0 \leq \frac{\theta_1}{\theta_2} \leq 1$ we have with $\rho = 1 - \frac{\theta_1}{\theta_2} \in [0, 1]$

$$\begin{aligned}
\dot{\varphi}_\theta &= \frac{-1}{\theta_2(\rho+1)}\left(0, \frac{\rho+1}{2}, \frac{\rho^2-1}{2}\right)^\top \\
&= \frac{-1}{\theta_2(\rho+1)}\left(\frac{1}{2}(-1,1,-1)^\top + \frac{1}{2}(1,\rho,\rho^2)^\top\right) \\
&= \frac{-1}{\theta_2(\rho+1)}\left(-\frac{1}{2}f(-1)^\top + \frac{1}{2}f(\rho)^\top\right).
\end{aligned}$$

Hence, the intersection of $\partial\mathcal{E}$ and the closure of $\{\mu\,\dot{\varphi}_\theta;\ \mu > 0,\ 0 \leq \frac{\theta_1}{\theta_2} \leq 1\}$ is $G_2 := \{(0, \frac{1+\rho}{2}, \frac{\rho^2-1}{2})^\top;\ \rho \in [0, 1]\}$.

For $-1 \leq \frac{\theta_1}{\theta_2} \leq 0$ we have with $\rho = 1 + \frac{\theta_1}{\theta_2} \in [0, 1]$

$$\dot{\varphi}_\theta = \frac{-1}{\theta_2(1+\rho)}\left(0, \frac{1+\rho}{2}, \frac{1-\rho^2}{2}\right)^\top$$

$$= \frac{-1}{\theta_2(1+\rho)} \left(\frac{1}{2}(1,1,1)^\top + \frac{1}{2}(-1,\rho,-\rho^2)^\top\right).$$

Hence, the intersection of $\partial\mathcal{E}$ and the closure of $\{\mu\,\dot\varphi_\theta;\ \mu > 0,\ 0 \le \frac{\theta_1}{\theta_2} \le 1\}$ is $G_3 := \{(0, \frac{1+\rho}{2}, \frac{1-\rho^2}{2})^\top;\ \rho \in [0,1]\}$.

For $\frac{\theta_1}{\theta_2} \le -1$ we have with $\rho = -\frac{\theta_2}{\theta_1} \in (0,1]$

$$\begin{aligned}
\dot\varphi_\theta &= \frac{\theta_1}{\theta_2^2}\left(0, \frac{-\theta_2}{2\theta_1}, \frac{1}{2}\right)^\top = \frac{\theta_1}{\theta_2^2}\left(0, \frac{\rho}{2}, \frac{1}{2}\right)^\top \\
&= \frac{\theta_1}{\theta_2^2}\left(\frac{1}{2}(-1,0,0)^\top + \frac{1}{2}\,\frac{1+\rho}{2}\,(1,1,1)^\top + \frac{1}{2}\,\frac{1-\rho}{2}\,(1,-1,1)^\top\right).
\end{aligned}$$

Hence, the intersection of $\partial\mathcal{E}$ and the closure of $\{\mu\,\dot\varphi_\theta;\ \mu > 0,\ \frac{\theta_1}{\theta_2} \le -1\}$ is $G_4 := \{(0, \frac{\rho}{2}, \frac{1}{2})^\top;\ \rho \in [0,1]\}$.

If $\overline{G} := G_1 \cup G_2 \cup G_3 \cup G_4$, then $G = \overline{G} \cup \{-c;\ c \in \overline{G}\}$.

A candidat for a maximin efficient design is $\delta_\alpha = \frac{\alpha}{2}e_{-1} + (1-\alpha)e_0 + \frac{\alpha}{2}e_1$ with $\alpha \in (0,1)$. To find the appropriate α one can determine the α so that $\arg\max_{c\in\overline{G}} c^\top \mathcal{I}(\delta_\alpha)^{-1} c$ has three elements:

At first note that

$$\mathcal{I}(\delta_\alpha) = \begin{pmatrix} 1 & 0 & \alpha \\ 0 & \alpha & 0 \\ \alpha & 0 & \alpha \end{pmatrix}$$

and

$$\mathcal{I}(\delta_\alpha)^{-1} = \frac{1}{\alpha\,(1-\alpha)} \begin{pmatrix} \alpha & 0 & -\alpha \\ 0 & 1-\alpha & 0 \\ -\alpha & 0 & 1 \end{pmatrix}.$$

If $c = (0, \frac{\rho}{2}, \frac{-1}{2})^\top \in G_1$, then

$$\begin{aligned}
& c^\top \mathcal{I}(\delta_\alpha)^{-1} c \\
&= \frac{1}{4\,\alpha\,(1-\alpha)}\,(\rho^2(1-\alpha)+1) \\
&\le \frac{2-\alpha}{4\,\alpha\,(1-\alpha)}
\end{aligned}$$

for all $\rho \in [0,1]$. In particular for $\alpha < \frac{2}{3}$ we have

$$c^\top \mathcal{I}(\delta_\alpha)^{-1} c < \frac{1}{\alpha}$$

for all $c \in G_1$. The same holds for $c \in G_4$.

If $c = (0, \frac{\rho+1}{2}, \frac{\rho^2-1}{2})^\top \in G_2$ or $c = (0, \frac{1+\rho}{2}, \frac{1-\rho^2}{2})^\top \in G_3$, then

$$\begin{aligned} & c^\top \mathcal{I}(\delta_\alpha)^{-1} c \\ = \quad & \frac{(\rho+1)^2(\rho^2 - 2\rho + 2 - \alpha)}{4\,\alpha\,(1-\alpha)}. \end{aligned}$$

Set $h_\alpha(\rho) := (\rho+1)^2(\rho^2 - 2\rho + 2 - \alpha)$. Then

$$h'_\alpha(\rho) = (\rho+1)(4\rho^2 - 4\rho + 2 - 2\alpha) = 0$$

and $\rho \in [0,1]$ if and only if $\rho = \frac{1}{2}(1 \pm \sqrt{2\alpha - 1})$. Because $h'_\alpha(0) = 2 - 2\alpha > 0$ and $h'_\alpha(1) = 4 - 4\alpha > 0$ the maximum points of $h_\alpha : [0,1] \to I\!R$ can be only $\rho_1 = 1$ and $\rho_2 = \frac{1}{2}(1 - \sqrt{2\alpha - 1})$. For these points we have

$$\frac{h_\alpha(\rho_1)}{4\,\alpha\,(1-\alpha)} = \frac{4\,(1-\alpha)}{4\,\alpha\,(1-\alpha)} = \frac{1}{\alpha}$$

and

$$\frac{h_\alpha(\rho_2)}{4\,\alpha\,(1-\alpha)} = \frac{-\alpha^2 - 8\alpha + 4\alpha\sqrt{2\alpha-1} - 2\sqrt{2\alpha-1} + 11}{16\,\alpha\,(1-\alpha)}.$$

If ρ_1 as well as ρ_2 should be the maximum points, i.e. $h_\alpha(\rho_1) = h_\alpha(\rho_2)$, then α should be a solution of

$$-\alpha^2 + 8\alpha + 4\alpha\sqrt{2\alpha-1} - 2\sqrt{2\alpha-1} - 5 = 0.$$

The solution of the equation is $\alpha^* = 23 - 10\sqrt{5} \approx 0.6393202$ because $\sqrt{2\alpha^* - 1} = \sqrt{5}\sqrt{(2-\sqrt{5})^2}$. Note that the numerical weights 0.31966, 0.36168, and 0.31966 for the design points -1, 0 and 1 of the maximin efficient design were already derived by Buonaccorsi and Iyer (1986) using a numerical search procedure.

Proposition 3.2 *The maximin efficient design for estimating the maximum point of a quadratic function with intercept in* $\Delta = \{\delta;\ supp(\delta) \subset [-1,1]\}$ *is* $\delta^* = \frac{\alpha^*}{2} e_{-1} + (1 - \alpha^*) e_0 + \frac{\alpha*}{2} e_1$ *with* $\alpha^* = 23 - 10\sqrt{5}$.

Proof. ¿From the above calculations we have that $c^\top \mathcal{I}(\delta^*)^{-1} c$ is maximal for $c = (0, \frac{\rho+1}{2}, \frac{\rho^2-1}{2})^\top$ and $c = (0, \frac{1+\rho}{2}, \frac{1-\rho^2}{2})^\top$ with $\rho = 1$ and $\rho = \frac{1}{2}(1 - \sqrt{2\alpha^* - 1}) = \sqrt{5} - 2$. I.e. we have $\{c_1, c_2, c_3\} = \arg\max_{c \in G} c^\top \mathcal{I}(\delta^*)^{-1} c$, where $c_1 = (0,1,0)^\top$, $c_2 = (0, \frac{\sqrt{5}-1}{2}, \frac{8-4\sqrt{5}}{2})^\top$, and $c_3 = (0, \frac{\sqrt{5}-1}{2}, \frac{4\sqrt{5}-8}{2})^\top$. This implies

$$c_1^\top \mathcal{I}(\delta^*)^{-1} f(t) = \frac{t}{23 - 10\sqrt{5}},$$

$$c_2^\top \mathcal{I}(\delta^*)^{-1} f(t) = \frac{-96 + 43\sqrt{5} - 8t\sqrt{5} + 18t + 2t^2 - t^2\sqrt{5}}{(-11 + 5\sqrt{5})\,(23 - 10\sqrt{5})},$$

$$c_3^\top \mathcal{I}(\delta^*)^{-1} f(t) = \frac{96 - 43\sqrt{5} - 8t\sqrt{5} + 18t - 2t^2 + t^2\sqrt{5}}{(-11 + 5\sqrt{5})\,(23 - 10\sqrt{5})},$$

and

$$c_1^\top \mathcal{I}(\delta^*)^{-1} c_1 = c_2^\top \mathcal{I}(\delta^*)^{-1} c_2 = c_3^\top \mathcal{I}(\delta^*)^{-1} c_3 = \frac{1}{23 - 10\sqrt{5}}.$$

Set $\lambda = \frac{2\sqrt{5}+22}{29}$. Then we have

$$\begin{aligned} &(1-\lambda)\,(c_1^\top \mathcal{I}(\delta^*)^{-1} f(t))^2 + \frac{\lambda}{2}\,(c_2^\top \mathcal{I}(\delta^*)^{-1} f(t))^2 \\ &\quad + \frac{\lambda}{2}\,(c_3^\top \mathcal{I}(\delta^*)^{-1} f(t))^2 - \frac{1}{23 - 10\sqrt{5}} \\ &= \; -\frac{2}{29}\,\frac{t^2(79 - 35\sqrt{5} - 79t^2 + 35t^2\sqrt{5})}{(23 - 10\sqrt{5})^2\,(-11 + 5\sqrt{5})^2} \\ &=: \; g(t). \end{aligned}$$

The first and the second derivative of $g(t)$ are

$$g'(t) = -\frac{4}{29}\,\frac{t(79 - 35\sqrt{5} - 158t^2 + 70t^2\sqrt{5})}{(23 - 10\sqrt{5})^2\,(-11 + 5\sqrt{5})^2}$$

and

$$g''(t) = -\frac{4}{29}\,\frac{79 - 35\sqrt{5} - 474t^2 + 210t^2\sqrt{5}}{(23 - 10\sqrt{5})^2\,(-11 + 5\sqrt{5})^2}.$$

Roots of g' are 0, $\frac{1}{2}\sqrt{2}$, and $-\frac{1}{2}\sqrt{2}$, where

$$g''(\frac{1}{2}\sqrt{2}) = g''(-\frac{1}{2}\sqrt{2}) = -\frac{8}{29}\,\frac{-79 + 35\sqrt{5}}{(23 - 10\sqrt{5})^2\,(-11 + 5\sqrt{5})^2} \approx 15.3.$$

Hence the maximum values of g are attained at -1, 0, and 1, where $g(-1) = g(0) = g(1) = 0$. According to Theorem 2.2 the design δ^* is maximin efficient for estimating the maximum point. $\square$

4 Maximin efficiency in nonlinear models

Assume a nonlinear model of the form $Y = \eta(t,\theta) + Z$. Estimating θ with the least squares estimator, the asymptotic covariance matrix has the form $\mathcal{I}_\theta(\delta)^{-1}$, where

$$\mathcal{I}_\theta(\delta) = \int f_\theta(t) f_\theta(t)^\top \delta(dt)$$

and $f_\theta(t) := \frac{\partial}{\partial\theta}\eta(t,\theta) \in I\!R^p$. In the following we use the D-criterion for defining maximin efficient designs, that is, we have $\Phi_\theta(\delta) = \det \mathcal{I}_\theta(\delta)^{-1}$. Locally optimal designs are

$$\delta_\theta \in \arg\min\{\det \mathcal{I}_\theta(\delta)^{-1};\ \delta \in \Delta\} = \arg\max\{\det \mathcal{I}_\theta(\delta);\ \delta \in \Delta\}.$$

Then maximin efficient designs δ^* are given by

$$\delta^* \in \arg\max\left\{\min_\theta \frac{\det \mathcal{I}_\theta(\delta_\theta)^{-1}}{\det \mathcal{I}_\theta(\delta)^{-1}};\ \delta \in \Delta\right\} = \arg\max\left\{\min_\theta \frac{\det \mathcal{I}_\theta(\delta)}{\det \mathcal{I}_\theta(\delta_\theta)};\ \delta \in \Delta\right\}$$

and are minimax designs for

$$\Psi_\theta(\delta) = \frac{\det \mathcal{I}_\theta(\delta_\theta)}{\det \mathcal{I}_\theta(\delta)}.$$

¿From the general Theorem 2.1 we obtain the following theorem for maximin efficiency in nonlinear models.

Theorem 4.1 *If $\mathcal{M} = \arg\min_{\theta\in\Theta} \frac{\det \mathcal{I}_\theta(\delta^*)}{\det \mathcal{I}_\theta(\delta_\theta)}$ is finite and Θ is compact, then δ^* is maximin efficient in Δ if and only if there exists $\lambda : \mathcal{M} \to [0,1]$ with $\sum_{\theta\in\mathcal{M}} \lambda(\theta) = 1$ and*

$$\sum_{\theta\in\mathcal{M}} \lambda(\theta)\, f_\theta(t)^\top \mathcal{I}_\theta(\delta^*)^{-1} f_\theta(t) \leq p$$

for all $t \in \mathcal{T}$.

Proof. Because the logarithm is strictly increasing we can work with

$$\Psi_\theta(\delta) = \ln\left(\frac{\det \mathcal{I}_\theta(\delta_\theta)}{\det \mathcal{I}_\theta(\delta)}\right) = \ln\det \mathcal{I}_\theta(\delta_\theta) - \ln\det \mathcal{I}_\theta(\delta)$$

instead of $\frac{\det \mathcal{I}_\theta(\delta_\theta)}{\det \mathcal{I}_\theta(\delta)}$. Then the continuity assumptions of Theorem 2.1 are satisfied and we have for the derivative of Ψ_θ at δ^*

$$\nabla_\theta(\delta^*, e_t) = p - f_\theta(t)^\top \mathcal{I}_\theta(\delta^*)^{-1} f_\theta(t)$$

(see e.g. Silvey 1980, page 44) so that Theorem 2.1 provides the assertion. □

4.1 One-parameter exponential growth curve

In the one-parameter exponential growth curve model we have $\eta(t,\theta) = e^{-\theta t}$, that is, $f_\theta(t) = -te^{-\theta t}$. Assume $\Theta = [0, K]$ with $K > 1$ and $\mathcal{T} = [0,1]$ so that $\Delta = \{\delta;\ \mathrm{supp}(\delta) \subset [0,1]\}$ is the set of designs. Hence, for every design $\delta \in \Delta$, the information matrix is

$$\mathcal{I}_\theta(\delta) = \int_0^1 t^2 e^{-2\theta t} \delta(dt).$$

The locally optimal designs are one-point designs. Even for many prior distribution, Bayesian designs are one-point designs (see Dette and Neugebauer 1997). However, a maximin efficient design cannot be a one-point design.

Proposition 4.1 *For sufficient large K a one-point design is not maximin efficient for the one-parameter exponential growth curve model.*

Proof. Let $\delta^* = e_\tau$ with $\tau \in [0,1]$ be an arbitrary one-point design. At first we are calculating the set

$$\mathcal{M} = \arg \min_{\theta \in [0,K]} \ln \left(\frac{\mathcal{I}_\theta(\delta^*)}{\mathcal{I}_\theta(\delta_\theta)} \right).$$

For $\theta \in [1,K]$ the locally optimal design is $\delta_\theta = e_{1/\theta}$ so that $\mathcal{I}_\theta(\delta_\theta) = \frac{1}{\theta^2} e^{-2}$ and

$$h_1(\theta) := \ln \left(\frac{\mathcal{I}_\theta(\delta^*)}{\mathcal{I}_\theta(\delta_\theta)} \right) = 2 \ln \tau - 2\theta\tau + 2 \ln \theta + 2.$$

For $\theta \in [0,1]$ the locally optimal design is $\delta_\theta = e_1$ so that $\mathcal{I}_\theta(\delta_\theta) = e^{-2\theta}$ and

$$h_2(\theta) := \ln \left(\frac{\mathcal{I}_\theta(\delta^*)}{\mathcal{I}_\theta(\delta_\theta)} \right) = 2 \ln \tau - 2\theta\tau + 2\theta = 2 \ln \tau + 2\theta(1-\tau).$$

The function h_1 is strictly increasing for $\theta \in [1, 1/\tau)$ and strictly decreasing for $\theta \in (1/\tau, K]$. The function h_2 is increasing with $h_2(1) = h_1(1)$. Because $h_1(k) \underset{(=)}{\geq} h_2(0)$ if and only if $\tau \underset{(=)}{\leq} \frac{1+\ln K}{K}$ we have

$$\begin{aligned} \mathcal{M} &= \{0\}, \quad \text{if} \quad \tau < \frac{1+\ln K}{K}, \\ \mathcal{M} &= \{0, K\}, \quad \text{if} \quad \tau = \frac{1+\ln K}{K}, \text{ and} \\ \mathcal{M} &= \{K\}, \quad \text{if} \quad \tau > \frac{1+\ln K}{K}. \end{aligned}$$

If $\tau < \frac{1+\ln K}{K}$, then $\delta^* = e_\tau$ is not locally optimal for $\theta = 0$ so that the necessary condition of Theorem 4.1 is not satisfied. Hence, in this case, $\delta^* = e_\tau$ is not maximin efficient. The same holds for $\tau > \frac{1+\ln K}{K}$.

For $\tau = \frac{1+\ln K}{K}$ the necessary condition of Theorem 4.1 says that a constant $\lambda \in [0,1]$ exists with

$$\begin{aligned} (4.6) \quad 1 &\geq \lambda f_0(t)^\top \mathcal{I}_0(e_\tau)^{-1} f_0(t) + (1-\lambda) f_K(t)^\top \mathcal{I}_K(e_\tau)^{-1} f_K(t) \\ &= \lambda\, t^2\, \tau^{-2} + (1-\lambda) t^2 e^{-2Kt} \tau^{-2} e^{2K\tau} \\ &= \lambda t^2 \frac{K^2}{(1+\ln K)^2} + (1-\lambda) t^2 \frac{K^2}{(1+\ln K)^2}\, e^{2+2\ln K - 2Kt} \end{aligned}$$

for all $t \in [0,1]$. In particular for $t = 1$, this means

$$1 \geq \lambda \frac{K^2}{(1+\ln K)^2} + (1-\lambda) \frac{K^2}{(1+\ln K)^2}\, e^{2+2\ln K - 2K}$$

which is equivalent to

$$(4.7)\quad \lambda \leq \frac{\frac{(1+\ln K)^2}{K^2} - e^{2+2\ln K-2K}}{1 - e^{2+2\ln K-2K}}.$$

For $t = 1/K$, the maximum value of $f_K(t)^\top \mathcal{I}_K(e_\tau)^{-1} f_K(t)$, the inequality (4.6) means

$$1 \geq \lambda \frac{1}{(1+\ln K)^2} + (1-\lambda)\frac{1}{(1+\ln K)^2} K^2$$

which is equivalent to

$$(4.8)\quad \lambda \geq \frac{(1+\ln K)^2 - K^2}{1-K^2}.$$

For large K the inequalities (4.7) and (4.8) cannot be satisfied simultanuously. This can be seen as follows: The inequalities (4.7) and (4.8) imply

$$\frac{(1+\ln K)^2 - K^2}{1-K^2} \leq \frac{(1+\ln K)^2 - K^2 e^{2+2\ln K-2K}}{K^2 - K^2 e^{2+2\ln K-2K}}$$

which is equivalent to

$$(4.9)\quad 2(1+\ln K)^2 K^2 - (1+\ln K)^2 K^4 e^{2(1-K)} - K^4 - (1+\ln K)^2 + K^4 e^{2(1-K)} \geq 0.$$

The leading term in the last inequality is $-K^4$ so that for large K the inequality is not satisfied. Hence, for large K there exists no $\lambda \in [0,1]$ so that condition (4.6) is satisfied for $t = 1$ and $t = 1/K$. This means that the necessary condition of Theorem 4.1 cannot be satisfied by a one-point design. $\square$

Numerical calculations showed that inequality (4.9) is already not satisfied for $K \geq 2.03$ so that inequality (4.6) is not satisfied for $t = 1$ and $t = 1/K$. For $1 < K < 2.03$ numerical simulations indicated that the necessary condition (4.6) is not satisfied for some $t \in (0,1)$ as well.

¿From the above results we now know that maximin efficient designs must have two or more support points. For practical purposes this is an advantage since the use of one-point designs is very restrictive. However, for calculating the maximin efficient design this is a disadvantage because we have to optimize over more than one parameter. Therefore, up to now we were not able to derive the maximin efficient design for this simple one-parameter exponential growth model.

Note that by a similar proof we also can show that one-point designs fail to be maximin efficient in inverse polynomial models of the form $\eta(t,\theta) = \frac{1}{1+\theta t}$ with $t \in [0,1]$. This can be related to the fact that the locally optimal designs for this model and for the exponential growth model are the same (see Atkinson and Bogacka 1997).

4.2 Two-parameter exponential growth curve

Using the results from the one-parameter exponential growth curve model we can show a similar result as for the one-parameter case. Namely, in the two-parameter case a two-point design with equal weights cannot be maximin efficient. For that we assume the two-parameter model given by $\eta(t,\theta) = \theta_1 e^{-\theta_2 t}$ so that $f_\theta(t) = e^{-\theta_2 t}(1, -\theta_1 t)^\top$,

$$\mathcal{I}_\theta(\delta) = \int e^{-2\theta_2 t} \begin{pmatrix} 1 & -\theta_1 t \\ -\theta_1 t & \theta_1^2 t^2 \end{pmatrix} \delta(dt)$$

and

$$\det \mathcal{I}_\theta(\delta) = \theta_1^2 \left[\int e^{-2\theta_2 t}\delta(dt) \int t^2 e^{-2\theta_2 t}\delta(dt) - \left(\int t e^{-2\theta_2 t}\delta(dt) \right)^2 \right].$$

We additionally assume that $\Delta = \{\delta;\ \mathrm{supp}(\delta) \subset [0,1]\}$ is the set of designs and that $\Theta = \{(\theta_1,\theta_2) \in I\!R^2;\ \theta_2 \in [0,K]\}$ with $K > 1$ is the set of unknown parameters.

In this setup locally optimal designs and even several Bayesian designs are of the form $\frac{1}{2}(e_0 + e_\tau)$, i.e. they are two-points designs with equal weights (see Dette and Neugebauer 1997). We are now going to show that, at least for large K, maximin efficient designs are not of this form.

At first note that $\frac{\det \mathcal{I}_\theta(\delta)}{\det \mathcal{I}_\theta(\delta_\theta)}$ is independent of θ_1 so that maximin efficieny concerns only θ_2. Hence, we can assume without loss of generality that $\theta_1 = 1$ and that only θ_2 is unknown. Consequently instead of $f_\theta(t)$ and $\mathcal{I}_\theta(\delta)$, we write $f_{\theta_2}(t)$ and $\mathcal{I}_{\theta_2}(\delta)$, respectively. We also use $\Theta_2 := [0,K]$ instead of Θ.

As in the one-parameter model we are firstly calculating the set

$$\mathcal{M} = \arg\min_{\theta_2 \in \Theta_2} \ln \left(\frac{\det \mathcal{I}_{\theta_2}(\delta^*)}{\det \mathcal{I}_{\theta_2}(\delta_{\theta_2})} \right)$$

for any design δ^* of the form $\delta^* = \frac{1}{2}(e_0 + e_\tau)$ with $\tau \in [0,1]$. For $\theta_2 \in [1,K]$ the locally optimal design is $\delta_{\theta_2} = \frac{1}{2}(e_0 + e_{1/\theta_2})$ so that $\det \mathcal{I}_{\theta_2}(\delta_{\theta_2}) = \frac{1}{4\theta_2^2} e^{-2}$ and

$$\ln \left(\frac{\det \mathcal{I}_{\theta_2}(\delta^*)}{\det \mathcal{I}_{\theta_2}(\delta_{\theta_2})} \right) = 2\ln\tau - 2\theta_2\tau + 2\ln\theta_2 + 2 = h_1(\theta_2).$$

For $\theta_2 \in [0,1]$ the locally optimal design is $\delta_{\theta_2} = \frac{1}{2}(e_0 + e_1)$ so that $\det \mathcal{I}_{\theta_2}(\delta_{\theta_2}) = \frac{1}{4} e^{-2\theta_2}$ and

$$\ln \left(\frac{\det \mathcal{I}_{\theta_2}(\delta^*)}{\det \mathcal{I}_{\theta_2}(\delta_{\theta_2})} \right) = 2\ln\tau + 2\theta_2(1-\tau) = h_2(\theta_2).$$

The functions h_1 and h_2 are defined as in the Section 4.1 for the one-parameter model. As in Section 4.1 we obtain

$$\mathcal{M} = \{0\}, \quad \text{if} \quad \tau < \frac{1 + \ln K}{K},$$

$$\mathcal{M} = \{0, K\}, \quad \text{if} \quad \tau = \frac{1 + \ln K}{K}, \text{ and}$$
$$\mathcal{M} = \{K\}, \quad \text{if} \quad \tau > \frac{1 + \ln K}{K}$$

and the fact that for $\tau \neq \frac{1+\ln K}{K}$ the design $\delta^* = \frac{1}{2}(e_0 + e_\tau)$ cannot be maximin efficient.

For $\tau = \frac{1+\ln K}{K}$ the necessary condition of Theorem 4.1 says that a constant $\lambda \in [0,1]$ exists with

$$(4.10) \quad 2 \;\geq\; \lambda f_0(t)^\top \mathcal{I}_0(\delta^*)^{-1} f_0(t) + (1-\lambda) f_K(t)^\top \mathcal{I}_K(\delta^*)^{-1} f_K(t)$$

for all $t \in [0,1]$. Because

$$\begin{aligned} f_{\theta_2}(t)^\top \mathcal{I}_{\theta_2}(\delta^*)^{-1} f_{\theta_2}(t) &= \frac{2\, e^{-2\theta_2 t}}{\tau^2} \left(\tau^2 - 2\tau t + t^2 + t^2 e^{2\theta_2 \tau}\right) \\ &= \frac{2\, e^{-2\theta_2 t}}{\tau^2} \left((\tau - t)^2 + t^2 e^{2\theta_2 \tau}\right) \end{aligned}$$

inequality (4.10) is equivalent to

$$2 \;\geq\; \lambda \frac{2}{\tau^2} \left((\tau - t)^2 + t^2\right) + (1-\lambda) \frac{2\, e^{-2Kt}}{\tau^2} \left((\tau - t)^2 + t^2 e^{2K\tau}\right)$$

which implies

$$\begin{aligned} (4.11) \quad 1 &\geq \lambda\, t^2\, \tau^{-2} + (1-\lambda) t^2 e^{-2Kt} \tau^{-2} e^{2K\tau} \\ &= \lambda t^2 \frac{K^2}{(1+\ln K)^2} + (1-\lambda) t^2 \frac{K^2}{(1+\ln K)^2}\, e^{2+2\ln K - 2Kt}. \end{aligned}$$

Inequality (4.11) coincides with inequality (4.6) of Section 4.1, and in Section 4.1 it was shown that for large K there exists no $\lambda \in [0,1]$ so that this inequality is satisfied for all $t \in [0,1]$. Hence, for $\delta^* = \frac{1}{2}(e_0 + e_\tau)$ the necessary condition of Theorem 4.1 cannot be fulfilled so that δ^* cannot be maximin efficient.

Acknowledgement: The second author thanks the Deutsche Forschungsgemeinschaft for supporting his short visits at the Free University of Berlin which allowed to write this paper.

References

[1] Atkinson, A.C. and Bogacka, B. (1997). Compound D- and D_s-optimum designs for determining the order of a chemical reaction. *Technometrics* **39**, 347-356.

[2] Buonaccorsi, J.P. and Iyer, H.K. (1986). Optimal designs for ratios of linear combinations in the general linear model. *J. Statist. Plann. Inference* **13**, 345-356.

[3] Dette, H. and Neugebauer, H.-M. (1997). Bayesian D-optimal designs for exponential regression models. *J. Statist. Plann. Inference* **60**, 331-349.

[4] Elfving, G. (1952). Optimum allocation in linear regression theory. *Ann. Math. Statist.* **23**, 255-262.

[5] Fedorov, V.V. (1980) Convex design theory. *Mathg. Operationsforsch. Statist., Ser. Statistics*, **11**, 403-413.

[6] Ford, I., Kitsos, C.P. and Titterington, D.M. (1989). Recent advances in nonlinear experimental design. *Technometrics* **31**, 49-60.

[7] McDonald, G.C. and Studden, W.J. (1990). Design aspects of regression-based ratio estimation. *Technometrics* **32**, 417-424.

[8] Müller, C.H. (1995). Maximin efficient designs for estimating nonlinear aspects in linear models. *J. Statist. Plann. Inference* **44**, 117-132.

[9] Pázman, A. (1986). *Foundations of Optimum Experimental Design.* Reidel, Dordrecht.

[10] Pukelsheim, F. (1993). *Optimal Design of Experiments.* John Wiley, New York.

[11] Sibson, R. (1972). Contribution to discussion of 'Results in the theory and construction of D-optimum experimental designs' by H.P. Wynn. *J. Roy. Statist. Soc., Ser. B* **34**, 181-183.

[12] Silvey, S.D. (1972). Contribution to discussion of 'Results in the theory and construction of D-optimum experimental designs' by H.P. Wynn. *J. Roy. Statist. Soc., Ser. B* **34**, 174-175.

[13] Silvey, S.D. (1980). *Optimal Design.* Chapman and Hall, London.

[14] Schervish, M.J. (1995). *Theory of Statistics.* Springer, New York.

[15] Schwabe, R. (1997). Maximin efficient designs. Another view at D-optimality. *Statistics & Probability Letters* **35**, 109-114.

Georg-August-University Göttingen
Institute of Mathematical Stochastics
Lotzestr. 13
D-37083 Göttingen
Germany
e-mail: chmuelle@math.uni-goettingen.de

Comenius University
Department of Probability
and Statistics, Mlynská dolina
SK-84215 Bratislava
Slovakia
e-mail: pazman@fmph.uniba.sk