On the functional approach to optimal designs for nonlinear models

Viatcheslav B. Melas
St. Petersburg State University

Abstract

This paper concerns locally optimal experimental designs for nonlinear regression models. It is based on the functional approach introduced in (Melas, 1978). In this approach locally optimal design points and weights are studied as implicitly given functions of the nonlinear parameters included in the model. Representing these functions in a Taylor series enables analytical solution of the optimal design problem for many nonlinear models. A wide class of such models is here introduced. It includes, in particular,three parameters logistic distribution, hyperexponential and rational models. For these models we construct the analytical solution and use it for studying the efficiency of locally optimal designs. As a criterion of optimality the well known D-criterion is considered.

Key words: nonlinear regression,experimental designs, locally optimal designs, functional approach,three parameters logistic distribution, hyperexponential models,rational models, D-criterion, implicit function theorem.

1991 AMS subject classification: 62K05 (primary), 62G02 (secondary).

1 Introduction

The modern optimal design theory [see monographs (Fedorov, 1972), (Silvey, 1980), (Pukelsheim, 1993) and collected papers (Kiefer, 1985)] relates mainly to linear (in parameters)regression models. This paper is devoted to the analytical study of locally optimal designs,introduced in (Chernoff, 1953), for nonlinear (in parameters) regression models. Up to now such designs have been investigated mainly for cases where these designs can be found in an explicit form or as a result of numerical procedures [see (Box, Lucas,

1959) for an early reference and (Han, Chaloner, 2003) and literature cited there for recent examples)].

A functional approach has been introduced by (Melas, 1978) for hyperexponential models. It is based on the study of optimal design points as implicitly given functions of values of parameters. Recently this approach was applied to rational models (Melas, 2001). Very convenient recurrent formulas for expanding implicit functions into Taylor series were introduced in (Dette, Melas and Pepelyshev, 2004). The present paper develops the functional approach for a wider class of nonlinear models. Note that the approach was developed for linear in parameters polynomial and trigonometrical models in a number of papers [(Melas, 2000), (Dette, Melas, 2002, 2003), (Dette, Melas and Pepelyshev (2002, 2004)].

The main results of the present paper consist of the following. First, for a wide class of models we prove that support points of locally D-optimal designs are real analytic functions of the nonlinear parameters.Second, we generalize the recurrent formulas, mentioned above, and demonstrate that the Taylor expansions allow to calculate the support points with a high precision. Third, we use the Taylor expansions to estimate the minimal efficiency of locally optimal designs if the nonlinear parameters vary in a given set. Also we show that the minimal efficiency can be substantially improved by an optimal choice of the initial values inside the set. In this way the locally optimal designs become very close to maximin efficient designs[(Müller, 1995),(Dette, Melas, Wong, 2004)]. A formal outline of the problem and a basic equation for the support points are given in Section 2. The main results are described in Sections 3–5. The proofs are somewhat lengthy and are differed to the Appendix.

2 Outline of the problem

2.1 Basic regression model

Let the experimental results $y_1, \ldots, y_N \in \mathbb{R}^1$be described by the equation

$$y_j = \eta(x_j, \Theta) + \varepsilon_j, \ j = 1, \ldots, N, \tag{1}$$

where $\eta(x, \Theta)$ is a function of a known form, $\Theta = (\theta_1, \ldots, \theta_m)^T$ is the vector of unknown parameters, $x_j \in \mathfrak{X}$, $\mathfrak{X}$ is a given set, $\{\varepsilon_j\}$ are random errors such that

$$E\varepsilon_j = 0, \ E\varepsilon_i\varepsilon_j = \delta_{ij}\sigma^2 h^2(x_i), \ i, j, = 1, \ldots, N, \tag{2}$$

2

δ_{ij} is the Kroneker symbol, σ^2 is unknown constant, and $h(x)$ is a known function, $h(x) > 0$, $x \in \mathfrak{X}$.

A discrete probability measure given by the table

$$\xi = \left(\begin{array}{ccc} x_1, & \ldots, & x_n \\ \omega_1, & \ldots, & \omega_n \end{array} \right),$$

where $x_i \in \mathfrak{X}$, $x_i \neq x_j$ $(i \neq j)$, $\omega_i > 0$, $\sum_{i=1}^{n}, \omega_i = 1$ will be called, as usual, an approximate experimental design.

Let approximately ω_i N experiments be performed at the point x_i, $i = 1, \ldots, n$ according to a design ξ, and now consider the least squares estimate $\Theta_{(N)}$ of the parameter vector Θ.

Let us denote by Θ_{tr} the true value of the parameter vector Θ appearing in our model (1)–(2). As it is known (Jennrich, 1969) under some regularity conditions $\Theta_{(N)} \to \Theta_{tr}$ almost surely and the vector $\sqrt{N}(\Theta_{(N)} - \Theta_{tr})$ has an asymptotically normal distribution with zero expectation and the variance matrix

$$\mathcal{D} = \sigma^2 M^{-1}(\xi, \Theta),$$

where $M(\xi, \Theta)$ is the information matrix,

$$M(\xi, \Theta) = \left(\sum_{l=1}^{n} f_i(x_l, \Theta) f_j(x_l, \Theta) \omega_l h^{-2}(x_l) \right)_{ij=1}^{m},$$

$f_i(x_l, \Theta) = \frac{\partial \eta(x_l, \Theta)}{\partial \theta_i}$, $\Theta = \Theta_{tr}$.

A design ξ^* maximizing the magnitude of $\det M(\xi, \Theta)$ under a fixed value $\Theta = \Theta^{(0)}$ in the class of all approximate designs will be called a locally optimal design.

Usually a locally optimal design depends only on a part of parameters (see Section 4). Without loss of generality assume that these parameters are $\theta_{m-r+1}, \ldots, \theta_m$ and denote $\Theta_1 = (\theta_1, \ldots, \theta_{m-r})^T$, $\Theta_2 = (\theta_{m-r+1}, \ldots, \theta_m)^T$. Our purpose is to study the dependence of ξ^* on Θ_2. Assuming that Θ_1 is fixed we will consider the matrix $M(\xi, \Theta_2) = M(\xi, \Theta)$.

2.2 The basic equation

In many practical problems $\mathfrak{X} = [a, b]$ and we will restrict our attention to this case.

Without loss of generality assume that

$$a \leq x_1 < x_2 < \ldots < x_n \leq b.$$

A triple (n_1, n_2, n_3), where $n_1(n_2)$ is the number of support points at the left (right) bound, $n_1, n_3 = 0$ or 1, $n_2 = n - n_1 - n_3$, will be called the type of design.

Let $\Theta_2 \in \Omega$, where Ω is a given open one-connected set of possible values of the parameters.

Consider the designs which are locally D-optimal in the class of designs with minimal support (that is with $n = m$). In this paper these designs will be called saturated locally D-optimal. Such designs often prove to be locally D-optimal in the class of all approximate designs.

It can be easily checked (Fedorov, 1972) that the weights in a saturated locally D-optimal design equal to $\omega_i = 1/m$, $m = 1, \ldots, m$.

Suppose that for any $z \in Z$ a saturated locally D-optimal design has a fixed type (n_1, n_2, n_3), $n_1 + n_2 + n_3 = m$. Consider the case $n_1 = 0$, $n_3 = 1$. In this case we will define a vector τ and the design ξ_τ in the following way

$$
\tau = (x_1, \ldots, x_{m-1})^T, \, a < x_1 < \ldots < x_{m-1} < b,
$$
$$
\xi_\tau = \begin{pmatrix} x_1 & \ldots & x_{m-1} & b \\ 1/m & \ldots & 1/m & 1/m \end{pmatrix}.
$$

All other cases can be considered in a similar way.

Let q be a given real analytic vector function on Ω such that

$$
\Theta_2 \to z = q(\Theta_2) \in R^r
$$

is a one-to-one correspondence and therefore the invert function $q^{-1}(z)$ is well defined at the set

$$
Z = \{z; \exists \Theta_2 \in \Omega, \, z = q(\Theta_2)\}.
$$

Define the following concept.

Definition 2.1 *The vector function*

$$
\tau^*(z) : Z \to V,
$$

where

$$
V = \left\{ tau = (x_1, \ldots, x_{m-1})^T; \, a < x_1 < \ldots < x_{m-1} < b \right\}
$$

will be called an optimal design function if for any fixed $z \in Z$ the design $\xi_{\tau^*(z)}$ is a saturated locally D-optimal design for $(\Theta^0)^T = (\Theta_1^T, \Theta_2^T(z))$, $\Theta_1(z) = q^{-1}(z)$.

4

Let us introduce an equation which implicitly determine an optimal design function.

Define the function

$$\varphi(\tau, z) = [\det M(\xi_\tau, \Theta(z)]^{1/m}, \tag{3}$$

where $\Theta^T(z) = \left(\Theta_1^T, \Theta_2^T(z)\right)$, $\Theta_2(z) = q^{-1}(z)$.

By the assumptions given above, for any fixed $z \in Z$ function $\varphi(\tau, z)$, $\tau \in V$ attains its maximum inside the set V. Therefore, vanishing of the derivatives

$$\frac{\partial}{\partial \tau_i} \varphi(\tau, z) = 0, \ i = 1, \ldots, m - 1.$$

is the necessary condition for equality $\tau = \tau^*(z)$ for any fixed $z \in Z$.

Set

$$g_i = g_i(\tau, z) = \frac{\partial}{\partial \tau_i} \varphi(\tau, z), \ i = 1, \ldots, m - 1, g = (g_1, \ldots, g_{m-1})^T.$$

Then

$$g(\tau, z) = 0 \tag{4}$$

at $\tau = \tau^*(z)$. This equation is called *the basic equation*. It enables reducing the problem to the analysis of implicit functions. Such analysis will be given in Section 5. And in the next section we will introduce assumptions providing that these functions are real analytic.

3 Analytic properties of the design functions

Let Ω, Z and $q : \Omega \to Z$ be such as described in the previous section.

Let $\Psi(z)$ be a polynomial, $\mathcal{N}$ be the set of roots of this polynomial,

$$\mathcal{N} = \{z \in R^k; \ \psi(z) = 0\}$$

and $\mathcal{N}$ be such that for any $\bar{z} \in \mathcal{N}$ there exists a sequence $z_{(k)}$, $k = 1, 2, \ldots$, $z_{(k)} \in Z$, $z_{(k)} \to \bar{z}$ with $k \to \infty$.

Remember that

$$f_i(x, \Theta) = \frac{\partial}{\partial \theta_i} \eta(x, \Theta), i = 1, \ldots, m,$$

$\Theta^T = (\Theta_1^T, \Theta_2^T)$, Θ_1 is a fixed vector and $\Theta_2 \in \Omega$.

Consider a class of regression functions $\eta(x, \Theta)$, $x \in [a, b]$, $\Theta \in R^m$ satisfying the following assumptions.

A1. *Functions*

$$f_i(x, \Theta)/h(x), \, i = 1, \ldots, m$$

are real analytic with $x \in [a, b]$, $\Theta_2 \in \Omega$.

A2. *With* $\Theta_2 \in \Omega$ *all saturated locally D-optimal designs have one and the same type* (n_1, n_2, n_3).

Let V be the set of all vectors $\tau = (\tau_1, \ldots, \tau_m)^T$ such that $a \leq \tau_1 \leq \ldots \leq \tau_m \leq b$ and no more than two coordinated are coincide.

A3.

$$\inf_{z \in Z \cap \mathcal{N}} \inf_{\tau \in V} \tilde{\varphi}(\tau, z) > 0,$$

where

$$\tilde{\varphi}(\tau, z) = \frac{\left[\det \left(f_i(\tau_j, \Theta(z))\right)_{i,j=1}^{m}\right]^2}{\Psi(z) \prod_{j > i}(\tau_j - \tau_i)^2}.$$

Note that with $\Psi(z) \equiv 1$, $\mathcal{N} = \emptyset$ this assumption means that with a fixed z the functions $f_i(x, \Theta(z))$, $i = 1, \ldots, m$ generate an extended Chebyshev system of the first order on $[a, b]$ [see (Karlin, Studden, 1966, Ch. 1)] for all $z \in Z$.

Let us co-define the function $\tilde{\varphi}(\tau, z)$ for $z \in \mathcal{N}$ by continuity (this is possible due to **A3**).

A4. *There exists* $z_{(0)} \in Z \cup \mathcal{N}$ *such that the equation system*

$$\frac{\partial}{\partial \tau_i} \tilde{\varphi}(\tau, z_{(0)}), \, i = n_1 + 1, \ldots, m - n_3$$

has a unique solution $\tau = \tau_{(0)}$.

Examples of functions η satisfying assumptions **A1**–**A4** will be given in the next section.

Now we can formulate our main analitical results.

Theorem 3.1 *Let assumptions* **A1**–**A4** *be satisfied. Then there exist a unique optimal design function* $\tau^*(z) : z \cup \mathcal{N} \to R^k$. *It is a real analytic vector function in* $Z \cup \mathcal{N}$ *and its Taylor coefficients can be calculated by recurrent formulas (16) given in Section 5.*

Note that Assumption **A4** is needed to secure the uniqueness of the optimal design function. It can be replaced by

A4'. *For any $z \in Z$ there exists a unique saturated locally D-optimal design for $\Theta^0 = \Theta(z)$.*

Remark 3.1 *The assertion of Theorem 3.1 remains true if the assumption **A4** is replaced by **A4'**.*

4   Examples

4.1   Three parameters logistic distribution

Consider the function

$$\eta(t, \alpha, \beta, \gamma) = \frac{\alpha e^{\gamma t + \beta}}{1 + e^{\gamma t + \beta}}.$$

It is called three parameters logistic distribution. By the substitution $x = e^t$, $\theta_1 = \alpha$, $\theta_2 = \gamma$, $\theta_3 = e^{-\beta}$ this function is reduced to

$$\eta(x, \Theta) = \frac{\theta_1 x^{\theta_2}}{\theta_3 + x^{\theta_2}}, \tag{5}$$

which is called the Hill equation in microbiological studies [see (Bezeau, Endrenyi, 1986)].

We will construct locally D-optimal designs for model (5) using the functional approach described above.

Assume that $x \in [a, b]$, $a \geq 0$, $\theta_1 \neq 0$, $\theta_3 > 0$.

By a direct calculation we receive

$$\det M(\xi, \Theta) = \theta_1^4 \theta_3^2 \det \bar{M}(\eta, \theta_3),$$

where

$$\begin{aligned}
\xi &= \left(\begin{array}{ccc} x_1 & x_2 & x_3 \\ 1/3 & 1/3 & 1/3 \end{array} \right), \; \eta_\xi = \left(\begin{array}{ccc} t_1 & t_2 & t_3 \\ 1/3 & 1/3 & 1/3 \end{array} \right), \\
t_i &= x_i^{\theta_2}, \; i = 1, 2, 3, \\
\bar{M}(\eta_\xi, \theta_3) &= \sum_{i=1}^{3} f(t_i, \theta_3) f^T(t_i, \theta_3), \\
f(t, \theta) &= \left(\frac{t}{\theta + t}, \frac{t}{(\theta + t)^2}, \frac{t \ln t}{(\theta + t)^2} \right)^T.
\end{aligned}$$

Table 4.1: Coefficients of the Taylor expansions for x_1 and x_2 in a vicinity of point $z = 0$

	0	1	2	3	4	5	6
x_1	0.15370	-0.09435	0.06747	-0.05117	0.04089	-0.03371	0.02845
x_2	0.61680	-0.20012	0.08251	-0.03885	0.02085	-0.01212	0.00754

Set
$$z = 1/\theta_3, \ r = 1, \ \Omega = [0, \infty), \ \psi(z) = z^6, \ \mathcal{N} = \{0\}. \tag{6}$$

Assumption **A1** follows here from the properties of elementary functions, **A2** and **A3** follows from the results of (Dette, Melas, Wong, 2003). It was also proved there that a locally D-optimal design has the type $(0, 2, 1)$ and is unique. This means that **A4'** holds for the considering model.

Thus due to Theorem 3.1 and Remark 3.2 it follows that support points of locally D-optimal designs are real analytic functions of z with $z \in [0, 1)$.

Let us consider the case $[a, b] = [0, 1], \theta_2 = 1$. For arbitrary $0 \le a < b, \theta_2$ optimal designs can be calculated by a scale transformation. With $\theta_3 \to \infty$, $z = \frac{1}{\theta_3} \to 0$ we receive

$$\frac{\det(f_i(x_j, \theta_3))}{z^6} \to \det \begin{pmatrix} x_1^2 & x_2^2 & 1 \\ x_1 & x_2 & 1 \\ x_1 \ln x_1 & x_2 \ln x_2 & 0 \end{pmatrix} := Q(x_1, x_2)$$

and
$$(x_1^*(z), x_2^*(z)) \to \arg \max_{0 < x_1 < x_2 < 1} Q(x_1, x_2).$$

Thus it is easy to calculate numerically that $x_1^*(0) = 0.1535$, $x_2^*(0) = 0.667$.

By the recurrent formulas (16) given in Section 5 we calculated the Taylor coefficients with $z_{(0)} = 0$. The first coefficients are represented in Table 4.1.

Let $\xi_{<n>}(z)$ be the design constructed by using n first coefficients and $\bar{z}_n$ is the maximal z such that

$$\max_{x \in [0,1]} \ |d(x, \xi_{<n>}(z)) - 3| \le 10^{-5}, \tag{7}$$
$$d(x, \xi) = f^T(x) M^{-1}(\xi, z) f(x),$$

8

Table 4.2: Coefficients of the Taylor expansions for x_1 and x_2 in a vicinity of point $z = 1$ by degrees of $(z - 1)$.

	0	1	2	3	4	5	6
x_1	0.09723	-0.03401	0.01308	-0.00530	0.00222	-0.00095	0.00041
x_2	0.47233	-0.10533	0.02743	-0.00791	0.00245	-0.00080	0.00027

Table 4.3: Coefficients of the Taylor expansions for x_1 and x_2 in a vicinity of point $z = 1$ by degrees of $(1/z - 1)$

	0	1	2	3	4	5	6
x_1	0.09723	0.03401	-0.02093	0.01314	-0.00844	0.00555	-0.00375
x_2	0.47233	0.10533	-0.07790	0.05838	-0.04431	0.03404	-0.02647

where

$$f(x) = \frac{\partial \eta(x, \Theta)}{\partial \theta_i}, \ M(\xi, z) := M(\xi, \Theta(z)),$$

$\Theta(z) = (1, 1, 1/z)^T$.

Numerical calculations show that $\bar{z}_{10} \approx 0.705$, $\bar{z}_{20} \approx 0.865$.

In a similar way we constructed expansions of the vector function $\tau^*(z) = (x_1^*(z), x_2^*(z))^T$ in a vicinity of point $z_{(0)} = 1$ by degrees of $(z - 1)$ and $(1/z - 1)$. The corresponding coefficients are presented in Tables 4.2 and 4.3, respectively. It proves that for the first expansion with twenty coefficients the inequality (7) holds with $0 < z \leq 2.7$. And for the second expansion with the same number of the coefficients ot holds for $0.6 \leq z \leq 13.8$.

The behavior of the design points for $0 \leq z \leq 10$ is presented at Fig.1. We used the first expansion for $z \leq 1$ and the second — for $1 \leq z \leq 10$ to construct the Fig. 1.

9

Table 4.4: Efficiency of designs ξ_0,ξ_1 and the points of locally D-optimal designs

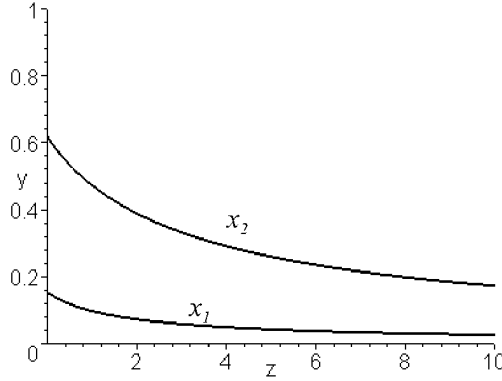| z	0.2	0.4	0.6	0.8	1.0
x_1	0.13690	0.12387	0.11333	0.10460	0.09723
x_2	0.57956	0.54751	0.51943	0.49456	0.47233
$\left(\frac{\det M(\xi_0,z)}{\det M(\xi_z,z)}\right)^{1/3}$	0.99343	0.97771	0.95681	0.93310	0.90801
$\left(\frac{\det M(\xi_1,z)}{\det M(\xi_z,z)}\right)^{1/3}$	0.94919	0.97468	0.98995	0.99774	1

Fig. 1. The dependence of the support points x_1 and x_2 on z

Note also that the efficiency of the limiting design (at the point $z_{(0)} = 0$) measured by the quantity

$$I(\xi, z) = \left(\frac{\det M(\xi, z)}{\det M\left(\xi_{\tau(z)}, z\right)}\right)^{1/3} , \ \xi = \xi_{\tau_{(0)}} := \xi(0)$$

proves to be very high with $z \leq 1$ ($\theta_3 \geq 1$). This efficiency is presented at Table 4.4.

At the same time the minimal efficiency of the design $\xi(1) = \xi_{\tau^*(1)}$ with $0 < z \leq 1$ is even more than that of $\xi(0) = \xi_{\tau^*(0)} = \xi_{\tau_{(0)}}$, see Table 4.4. Moreover, numerical calculations show that the design $\xi(z^*) = \xi_{\tau^*(z^*)}$ with $z^* = 0.5$ have maximum of the minimal efficiency at the interval $(0, 1]$ among locally D-optimal designs at points $z = 0.1, \ldots, 0.9, 1$. Its minimal efficiency is equal to 0.981.

Note that a maximin efficient D-optimal design that a the design maximizing the minimum by $z \in [0.1, 1]$ of the efficiency among all (approximate) designs,was constructed numerically in (Dette, Melas, Wong, 2004). This design is very close to $\xi(0.5)$ and has the minimal efficiency 0.982.

A similar calculation was performed for the interval$[1,10]$ for z. It showed that the design $\xi(4)$, the best design among $\xi(1)$, $\xi(2)$,. . .,$\xi(10)$,has minimal efficiency 0.8407. The maximin efficient design calculated in(Dette, Melas, Wong, 2004) has four support points with unequal weights and its minimal efficiency equals 0.885.But, for example, design $\xi(1)$, locally optimal design for z=1, has the minimal efficiency 0.5430 on $[1,10]$. This design is rather bad! It requires almost twice more observations than $\xi(4)$ to achieve the same accuracy of the parameters if true value of z equal to 10.

Thus we see that the approach allows very efficient calculation of locally D-optimal design and secure a study of their efficiency.

We conclude also that locally D-optimal designs could be very efficient if the initial values are chosen in the optimal way inside given intervals of possible values

4.2   Hyperexponential models

Let

$$\eta(x,\Theta) = \sum_{i=1}^{k} \theta_i e^{-\theta_{i+k} x},$$

$x \in \mathfrak{X} = [0, d]$, where d is sufficiently large and $\theta_i \neq 0$, $\theta_{i+k} > 0$, $i = 1, \ldots, k$, $\theta_{i+k} \neq \theta_{j+k}$ $(i \neq j)$, and $h(x) \equiv 1$.

Functions of such a form generate an important class of solutions of linear differential equations, which often occur in practice.

Consider the problem of finding designs to be locally optimal in the class of approximate designs with the number of points equal to the number of parameters. It can be verified that with $k = 1, 2$ such designs are locally optimal among all approximate designs [see (Dette, Melas, Wong, 2003)].

For $k = 1$ an immediate calculation shows that $x_1^* = 0$, $x_2^* = 1/\theta_2$.

Let $k \geq 2$, $z_i = 1 - \theta_{i+k}$, $i = 1, \ldots, k - 1$,

$$
\begin{aligned}
z_k &= \sum_{i=1}^{k} \theta_{i+k}/k = 1, \ z = (z_1, \ldots, z_{k-1}), \ z_{(0)} = (0, \ldots, 0), \\
\psi(z) &= \prod_{1 \leq i < j \leq k} (\theta_{i+k} - \theta_{j+k})^4, \ Z = (-1, 1)^{k-1}.
\end{aligned}
$$

11

Assumption **A1** is obviously true here and **A2–A4** follow from results of (Melas, 1978). Moreover the following proposition was proved there.

Proposition 4.1 *Under the above conditions there exists the unique solution of equation 4) with $z = z_{(0)}$.*

$$\tau_{(0)} = (\gamma_1/2, \ldots, \gamma_{2k-1}/2),$$

where $\gamma_1, \ldots, \gamma_{2k-1}$ are zeros of $L_{2k-1}^1(x)$, the Laguerre polynomial of degree $2k - 1$ with the parameter equal 1.

Due to Theorem 3.1 it follows that optimal design-function is uniquely determined and is analytical real vector function for $z \in Z$.

This assertion was proved in the paper cited in another sequence of arguments. However, in that paper the expansion of optimal design-function into Taylor series was not performed. This can be done with the help of formulas from Section 5.

Let $k = 2$,

$$\begin{aligned}
\eta(x, \Theta) &= \theta_1 e^{-\theta_3 x} + \theta_2 e^{-\theta_4 x}, \theta_1, \ \theta_2 \neq 0, \ \theta_3, \theta_4 > 0, \ \theta_3 \neq \theta_4, \\
\mathfrak{X} &= [0, \infty), \ (\theta_3 + \theta_4)/2 = 1.
\end{aligned}$$

Let us build the expansion of the vector function

$$\tau(z) = (x_2^*(z), x_3^*(z)), x_4^*(z)), \ z = (\theta_3 - \theta_4)/2$$

into series by degrees of z in a vicinity of the point $z = 0$.

Since $\varphi(\tau, z) = \varphi(\tau, -z)$ all odd coefficients are zeros.

Denote

$$\tau_{<2i>}(z) = \tau_{(0)} + \sum_{t=1}^{i} \tau_{(2t)} z^{2t}, \ i = 1, 2, \ldots$$

The coefficients for $t = 0, 1, \ldots, 6$ are presented in Table 4.5.

Let us consider the following problem: how many coefficients should be used in order to calculate locally optimal designs enough accurately? Numerical calculations allow to estimate the efficiency of designs $\xi_{\tau_{<2t>}(z)}$. These efficiences are shown at Table 4.6 for $t = 0, 1, \ldots, 9$.

Table 4.5: Coefficients τ_{2t}, $t = 0, 1, \ldots, 6$

0	1	2	3	4	5	6
0.46791	0.02919	0.00305	0.00056	0.00022	0.00008	-0.00005
1.65270	0.36419	0.21113	0.15971	0.13371	0.11650	0.10252
3.87938	2.00661	1.86581	1.92887	2.04481	2.16523	2.26335

12

Table 4.6: The efficiency of designs $\xi_{\tau_{<t>}(z)}$

$z\backslash t$	0	1	2	3	4	5	6	7	8	9
0.50	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.70	0.90	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.80	0.80	0.93	0.97	0.99	1.00	1.00	1.00	1.00	1.00	1.00
0.85	0.72	0.88	0.94	0.97	0.98	0.99	1.00	1.00	1.00	1.00
0.90	0.61	0.79	0.87	0.92	0.95	0.97	0.98	0.99	0.99	1.00
0.95	0.45	0.61	0.71	0.78	0.83	0.87	0.90	0.93	0.94	0.96
0.97	0.35	0.49	0.58	0.65	0.71	0.76	0.80	0.84	0.86	0.89

We can conclude from this table that if $0 < z \leq 0.7$ we need only one or two nonzero coefficients! But for $z = 0.9$ we need twenty coefficients (half of which are zeros). The table shows also that with $0 < z \leq 0.9$ the expansion allows to calculate locally optimal designs with a high precision.

Consider now the problem of the optimal choice of the initial values inside a given set. As in the previous section we will find a design maximizing the minimum of the efficiency $I(\xi, z)$ among locally optimal designs. Let us consider the case $0.1 \leq z \leq 0.9$. In this case the design $\xi(0.7)$ proves to be the best among the designs $\xi(0.1), \ldots, \xi(0.9)$. This design has the minimal efficiency 0.80768. Table 4.7 shows efficiencies of this design for different values of z. It presents also supprt points of locally D-optimal designs.

4.3 Rational regression

Consider the regression function of the form

$$\eta(x, \Theta) = \sum_{i=1}^{k} \theta_i / (x + \theta_{i+k}), \qquad (8)$$

$x \in \mathfrak{X} = [0, d]$, where d is sufficiently large, $\theta_i \neq 0$, $\theta_{i+k} > 0$, $i = 1, \ldots, k$, $\theta_{i+k} \neq \theta_{j+k}$, $(i \neq j)$, and $h(x) \equiv 1$.

Typically, models are chosen as approximations to unknown functions by a linear combination of functions of a certain type. Rational approximations can result in the models with fewer unknown parameters than the more habitual polynomial models [see (Petrushev, Popov, 1987)]. Some results on locally D-optimal designs for rational models were obtained in (He et al., 1996). Let z, $z_{(0)}$, $\psi(z)$ and Z be the same as in Section 3.1. The following results were announced in (Melas, 2001) without a proof.

13

Table 4.7: Efficiency of design $\xi(0.7)$ and the points of locally D-optimal designs

z	0.1	0.2	0.3	0.4	0.5
x_1	0.46820	0.46908	0.46056	0.47266	0.47541
x_2	1.65635	1.66762	1.68732	1.71714	1.76011
x_3	3.89941	3.96276	4.07665	4.25772	4.53863
$\left(\dfrac{\det M(\xi_{0.7},z)}{\det M(\xi_z,z)}\right)^{1/4}$	0.81739	0.83572	0.86493	0.90259	0.94428

z	0.6	0.7	0.8	0.9
x_1	0.47885	0.48303	0.48801	0.49379
x_2	1.82190	1.91409	2.06459	2.36561
x_3	4.98876	5.77821	7.43519	12.57015
$\left(\dfrac{\det M(\xi_{0.7},z)}{\det M(\xi_z,z)}\right)^{1/4}$	0.98188	1	0.96706	0.80768

Proposition 4.2 *For regression function (8) and for any fixed z in Z there exists a unique locally D-optimal design. It consists of m points, one of which is zero. Assumptions* **A1**–**A4** *are satisfied.*

Proposition 4.3 *The vector $\tau_{(0)}$ consists of zeros of the polynomial $q(x)$ of $2k - 1$ degree, where $q(x)$ is the unique solution of the differential equation (22) given in Section 6.*

Due to Theorem 3.1 it follows that the optimal design-function is a real analytical vector function.

For $k = 1$ it can be verified by an immediate calculation that the locally optimal design is $\{0, 1/\theta_2; 1/2, 1/2\}$.

For $k = 2$ the optimal design-function can be found in an explicit form as well. By the equation (22) given in Section 6 we can verify that

$$
x_3^* = \sqrt{\theta_3\theta_4}, \ x_{2,4}^* = \frac{\sqrt{\theta_3\theta_4}}{2}\left(-\Delta/2 - 1 \pm \sqrt{(\Delta/2 + 1)^2 - 4}\right),
$$
$$
\Delta = -(\theta_3 + \theta_4) - 3 - \sqrt{(\theta_3 + \theta_4 + 3)^2 + 24}.
$$

For $k = 3$ let us construct the expansion of the optimal design-function into a Taylor series.

14

Let $k = 3$,

$$\eta(x, \Theta) = \frac{\theta_1}{x + \theta_4} + \frac{\theta_2}{x + \theta_5} + \frac{\theta_3}{x + \theta_6}, \ \theta_1, \theta_2, \theta_3 \neq 0,$$

$\theta_4 > \theta_5 > \theta_6 > 0$, $\mathfrak{X} = [0, \infty)$, $(\theta_4 + \theta_5 + \theta_6)/3 = 1$.

Using proposition 4.3 we find $\tau_{(0)} \approx (0.09, 0.36, 1, 2.76, 10.78)$.

Set $u = (1 - \theta_4)(1 - \theta_5)$, $v = (2 - \theta_4 - \theta_5)$. Note that the points $x_{i+1}^*(\theta_4, \theta_5, \theta_6)$, $i = 1, \ldots, 5$ can be represented as

$$\tau_i(u, v) = \sum_{s_1=0}^{\infty} \sum_{s_2=0}^{\infty} \tau_{i(s_1, s_2)} u^{s_1} v^{s_2}. \tag{9}$$

The coefficients of this expansion are represented in Table 4.8.

Table 4.8: Coefficients for the rational model, $k = 3$

0, 0	0, 1	1, 0	0, 2	1, 1	2, 0
0.0928	0.0449	0	-0.0155	0.0540	-0.0449
0.3616	0.1514	0	-0.0481	0.1577	-0.1514
1.0000	0.3333	0	-0.0955	0.2864	-0.3333
2.7654	0.6861	0	-0.1803	0.4991	-0.6861
10.7802	1.9661	0	-0.5087	1.3892	-1.9661

Let $\tau_{<i>} = \tau_{<i>}(u, v)$ be the segment of the series (9) containing coefficients with $s_1 + s_2 \leq i$, $i = 1, 2, \ldots$. The efficiency of designs received from $\tau_{(0)}$, $\tau_{<i>}$, $i = 1, \ldots, 6$ by adding the point $x_1^* = 0$ is shown in Table 4.9.

Table 4.9: The efficiency of designs $\xi_{\tau_{<i>}}$, $i = 0, \ldots, 6$

$z_1 + z_2$	z_1	z_2	0	1	2	3	4	5	6
0.5	0.2	0.3	1.00	0.99	1.00	1.00	1.00	1.00	1.00
0.5	-0.2	0.7	0.96	0.98	1.00	1.00	1.00	1.00	1.00
0.7	0.3	0.4	0.99	0.98	1.00	1.00	1.00	1.00	1.00
0.7	0.0	0.7	0.96	0.96	1.00	1.00	1.00	1.00	1.00
0.9	0.4	0.5	0.97	0.94	1.00	1.00	1.00	1.00	1.00
0.9	0.1	0.8	0.90	0.88	0.99	0.99	1.00	1.00	1.00
1.1	0.5	0.6	0.93	0.88	1.00	0.98	1.00	1.00	1.00
1.1	0.3	0.8	0.87	0.82	0.99	0.97	1.00	0.99	1.00

5 The study of the basic equation

Let us study the equation (4) for the vector function $g(\tau, z)$ of a general form.

5.1 Smoothness and analycity of implicit functions

Assume that m and r are arbitrary natural numbers, $m \geq 2$ and that V and Z are bounded subsets of $\mathbb{R}^{m-1}$ and $\mathbb{R}^r$, respectively. We also assume that V and Z are simply connected sets.

Let $\varphi(\tau, z)$, $\tau \in V$, $z \in Z$ be a function of a general form possessing continuously differentiable (by τ) derivatives $g_i(\tau, r) = \frac{\partial}{\partial \tau_i} \varphi(\tau, z)$, $i = 1, \ldots, m-1$, $g(\tau, z) = (g_1(\tau, z), \ldots, g_{m-1}(\tau, z))^T$,

$$J(\tau, z) = \left(\frac{\partial^2}{\partial \tau_i \partial \tau_j} \varphi(\tau, z) \right)_{i,j=1}^{m-1}.$$

Consider the case when $\varphi(\tau, z) \equiv 0$ at some fixed points $z \in Z$. Let $\mathcal{N}$ be the set of all such points. Assume that there exists an algebraic polynomial $\psi(z)$ such that $\psi(z) = 0$ with $z \in \mathcal{N}$, $\psi(z) \neq 0$, $z \bar{\in} \mathcal{N}$ and the function

$$\tilde{\varphi}(\tau, z) = \varphi(\tau, z) / \psi(z)$$

can be extended for points $z \in \mathcal{N}$ in such a way that $\tilde{\varphi}(\tau, z)$ is twice continuously differentiable by τ with $(\tau, z) \in V \otimes Z$. Denote

$$\tilde{g}(\tau, z) = g(\tau, z) / \psi(z), \quad \tilde{J}(\tau, z) = J(\tau, z) / \psi(z).$$

Let us study the equation

$$\tilde{g}(\tau, z) = 0, \tag{10}$$

where $z \in Z$, $\tau \in V$.

Consider a point $z_{(0)} \in Z$ and the following condition.

(a) Equation (10) with $z = z_{(0)}$, $\tau \in V$ possesses a unique solution $\tau = \tau_{(0)}$, while

$$\det \tilde{J}(\tau_{(0)}, z_{(0)}) \neq 0.$$

Let U be a vicinity of the point $(\tau_{(0)}, z_{(0)})$, $U \subset V \otimes Z$.

Let us introduce the following conditions. Assume $K \geq 2$ to be a natural number.

16

(b) The vector function $\tilde{g}(\tau, z)$ is K times continuously differentiable in $(\tau, z) \in U$.

(c) The vector function $\tilde{g}(\tau, z)$ is a real analytical vector function for $(\tau, z) \in U$.

Theorem 5.1 *Assuming that conditions (a) and (b) are fulfilled. Then in some vicinity $Z_{(0)}$ of the point $z_{(0)}$ there exists a unique vector function $\tau(z)$: $Z \to V$ such that the following relations hold: $\tau(z) \in V$ and $\tilde{g}(\tau(z), z) = 0$. This vector function is $K - 1$ times continuously differentiable and satisfies the equations*

$$\tilde{J}\big(\tau(z), z\big)\tau'_{z_i}(z) = \big(\tilde{g}(\tau, z)\big)'_{z_i}\big|_{\tau = \tau(z)}, \ i = 1, 2, \ldots, r.$$

If the condition (c) is also fulfilled then $\tau(z)$ is a real analytical vector function with $z \in Z_0$. If $U = V \otimes Z$ and the condition (a) is fulfilled for any $z_{(0)} \in Z$ then the above assertion holds with $Z_{(0)} = Z$.

Note that the theorem is an obvious corollary of the Implicit Function Theorem (Gunning, Rossi, 1965, Ch. 1).

Since determinant of a matrix is an algebraic sum of some multiplications of its elements conditions (b) and (c) will be satisfied if the functions $f_i(x, z)$ ($i = 1, \ldots, m$) are K times differentiable and are real analytical functions, respectively, by the collection of variables $x, z_1, \ldots, z_r$.

To secure the verification of the condition (a) we will elaborate a representation of the Jacobi matrix.

5.2   Jacobian of the basic equation

First we analyze the Jacobian of the basic equation for functions $\varphi(\tau, z)$ of the general kind that can be represented as the minimum of some convex function.

Let m, k, t be arbitrary real numbers, $T \subset \mathbb{R}^{m-1}, Z \subset \mathbb{R}^k, \mathfrak{A} \subset \mathbb{R}^t$ — arbitrary open sets, where $\mathfrak{A}$ is convex.

Consider the function $q(\tau, a, z), \tau \in T, a \in \mathfrak{A}, z \in Z$ that satisfies the following conditions: function $q(\tau, a, z)$ is twice continuously differentiable along τ and a; function $q(\tau, a, z)$ is strictly convex along a.

Moreover, let function $\varphi(\tau, z)$ have the form

$$\varphi(\tau, z) = \min_{a \in \mathfrak{A}} q(\tau, a, z), \tag{11}$$

17

where the minimum is attained for any $\tau \in T, z \in Z$. Since the function $q(\tau, a, z)$ is convex along a, this minimum is attained on the unique vector $a = \tilde{a} = \tilde{a}(\tau, z)$. Therefore, function $\varphi(\tau, z)$ is twice continuously differentiable along τ.

Let for any fixed z there exists a point $\tilde{\tau} = \tilde{\tau}(z)$ satisfied the equation $\frac{\partial}{\partial \tau} \varphi(\tau, z) = 0$.

Consider the following matrices

$$E = \left(\frac{\partial^2}{\partial \tau_j \partial \tau_i} q(\tau, a, z) \right)_{i,j=1}^{m-1},$$

$$B = \left(\frac{\partial^2}{\partial \tau_j \partial a_i} q(\tau, a, z) \right)_{i,j=1}^{t,m-1}, \tag{12}$$

$$D = \left(\frac{\partial^2}{\partial a_j \partial a_i} q(\tau, a, z) \right)_{i,j=1}^{t}$$

at $\tau = \tilde{\tau}$, $a = \tilde{a}(\tilde{\tau}, z)$. It follows from above conditions that matrix D is positive definite and hence the inverse matrix D^{-1} exists.

Theorem 5.2 *Under the above conditions the following formula is valid:*

$$J(\tilde{\tau}(z), z) = E - B^T D^{-1} B.$$

Apply this theorem to function $\varphi(\tau, z)$, defined by formula (3).

Denote the set of all positive defined $m \times m$ matrices $A = (a_{ij})$, such that $a_{mm} = 1$ by $\mathcal{A}$.

Assign a number $\nu = \nu(i, j)$ to each pair of indices (i, j), $i \leq j$, $i, j = 1, \ldots, m$, $(i, j) \neq (m, m)$. For any vector $a \in \mathbb{R}^t$ define a matrix $A(a)$ that satisfies the following relations

$$a_{ji} = a_{ij} = a_{\nu(i,j)}, \ a_{mm} = 1, \ i, j = 1, \ldots, m, \ i \leq j.$$

Define set $\mathfrak{A}$ as

$$\mathfrak{A} = \{ a \in \mathbb{R}^t \, : \, A(a) \in \mathcal{A} \}.$$

Evidently, $\mathfrak{A}$ is open and convex in $\mathbb{R}^t$. Introduce the function

$$q(\tau, a, z) = (\det A(a))^{-1/m} tr \, (A(a) M(\xi, z)) / m. \tag{13}$$

Consider the function $\varphi(\tau, z) = (\det M(\xi, z))^{1/m}$. It is known (Karlin, Studden, 1966, Ch. 10.2) that formula (12) is valid for this function. It can

18

also be checked that the function (13) possesses the required properties. Therefore, by Theorem 5.2

$$J(\tau^*(z), z) = E - B^T D^{-1} B. \tag{14}$$

Set $\delta(a) = (\det A(a))^{-1/m}$. It is easy to verify by direct differentiation that the following formulas are valid for matrices B and E.

$$
\begin{aligned}
E &= diag\{E_{11}, \ldots, E_{m-1\,m-1}\}, \\
E_{ii} &= \delta(a^*) \left.\frac{\partial^2}{\partial x^2}(f^T(x)A(a^*)f(x))\right|_{x=x_{i+1}^*}, i = 1, \ldots, m-1, \\
A(a^*) &= const\,(M(\xi^*, z))^{-1}, \\
B &= (b_{\nu k})_{\nu, k=1}^{t, m-1}, \\
b_{\nu k} &= 2\delta(a^*) \left.\frac{\partial}{\partial x}(f_i(x)f_j(x))\right|_{x=x_k^*}, \nu = \nu(i, j).
\end{aligned}
\tag{15}
$$

Note that the matrix $J = J(\tau^*(z), z)$ is negative definite and hence nonsingular provided at least one of the following conditions is satisfied:

1) all diagonal elements of matrix E are negative;

2) matrix B is of full rank.

Indeed, matrix $B^T D^{-1} B$ has the form SS^T, hence, it is nonnegative definite in the general case and positive definite if matrix B has full rank. Since $J = E - B^T D^{-1} B$, then J is negative definite if either of conditions 1)-2) is valid.

5.3 On the representation of implicit functions as Taylor series

It is well known that derivatives of implicit functions can be calculated with the help of indefinite coefficients techniques, as introduced by Euler. In this section we offer recurrent formulas convenient for the implementation in software packages such as Maple and Mathcad. These formulas are a generalization for the multidimensional case of formulas introduced in (Dette, Melas, Pepelyshev, 2004).

Let us assume that $s = (s_1, \ldots, s_r)$, where $s_i \geq 0$, $i = 1, \ldots, r$ are integers. For an arbitrary (scalar, vector or matrix) function $\mathcal{F}$ denote

$$(\mathcal{F}(z))_{(s)} = \frac{1}{s_1! \ldots s_r!} \frac{\partial^{s_1}}{\partial z_1^{s_1}} \cdots \frac{\partial^{s_r}}{\partial z_r^{s_r}} \mathcal{F}(z)|_{z=z_{(0)}}.$$

Introduce also the notation

$$S_t = \{s = (s_1, \ldots, s_r); s_i \geq 0, \sum_{i=1}^{r} s_i = t\},$$

19

$t = 0, 1, \ldots,$

$$(z - z_{(0)})^s = (z_1 - z_{1(0)})^{s_1} \ldots (z_r - z_{r(0)})^{s_r}.$$

Let the function $\psi(z)$ be of the form

$$\psi(z) = (z - z_{(0)})^l \bar{\psi}(z),$$

where $l = (l_1, \ldots, l_r), l_i \geq 0$, $i = 1, \ldots, r$ are integers, $\bar{\psi}(z)$ is a homogeneous polynomial of degree $p \geq 0$,

$$\bar{\psi}(z) = \sum_{s \in S_p} a_{(s)}(z - z_{(0)})^s,$$

such that $a_{(p,0,\ldots,0)} \neq 0$.

Let

$$
\begin{aligned}
I_t &= U^t_{j=0} S_j, \\
\tau_{<I_t>}(z) &= \sum_{s \in I_t} \tau_{(s)}(z - z_{(0)})^s, \quad \tau_{(s)} = (\tau(z))_{(s)}, \\
J_{(l)} &= \left(J(\tau_{(0)}, z) \right)_{(l)}.
\end{aligned}
$$

At first let $p = 0$. Note that under condition (a) the matrices $J_{(s)}, s_i \leq l_i$, $i = 1, \ldots, r$, $s \neq l$ are zero matrices and $\det J_{(l)} \neq 0$.

Theorem 5.3 *Under the conditions (a) and (b) for the function $\tau(z)$, defined in Theorem 5.1 the following formulas take place*

$$(\tau(z))_{(s)} = -J^{-1}_{(l)} g(\tau_{<I>}(z), z)_{(s+l)} \tag{16}$$

where $I = I_{t-1}$, $s \in S_t$, $t = 1, 2, \ldots, K - 1$.

If the condition (c) is also fulfilled then these formulas hold for $t = 1, 2, \ldots$.

Thus, if $\tau_{(0)}$ is known, coefficients $\{\tau_{(s)}\}$ can be calculated in the following way. At the step t $(t = 1, 2, \ldots)$ calculate all coefficients with indices from S_t by formula (16). This calculation can be easily performed by a computer with the help of packages such as Maple or Mathcad.

Consider now the case $p > 0$.

Define the set

$$\hat{S}_t = \{ s = (s_1, \ldots, s_r); s_i \geq 0, i = 1, \ldots, r, s_1 + 2 \sum_{i=2}^r s_i = t \}.$$

20

Let

$$\hat{I}_t = U_{j=0}^t \hat{S}_j, \quad q = (p, 0, \ldots, 0),$$

$$J_{(l+q)} = \left(J(\tau_{(0)}, z) \right)_{(l+q)}.$$

It can be verified that, under condition (a), $\det J_{(l+q)} \neq 0$.

Theorem 5.4 *With $p > 0$ Proposition 5.3 remains true with formula (16) replaced by the formula*

$$\left(\tau(z) \right)_{(s)} = -J_{(l+q)}^{-1} g\left(\tau_{<I>}(z), z \right)_{(s+l+q)},$$

where $s \in \hat{S}_t$, $I = \hat{I}_{t-1}$, $t = 1, 2, \ldots$.

Note that q can be replaced by any vector of the form $(0, \ldots, p, 0, \ldots, 0)$.

6   Appendix: Proofs

We begin with the proofs for propositions of the last section.

Proof of Theorem 5.2. Due to the necessary condition for an extremum point we have

$$\frac{\partial}{\partial a} q(\tau, a, z) = 0$$

with an arbitrary fixed $z \in Z$ and with $\tau = \tilde{\tau} = \tilde{\tau}(z)$, $a = \tilde{a} = \tilde{a}(z, \tilde{\tau}(z))$.

Consider this vector equality at fixed z and arbitrary a and τ as an equations system which implicitly defines a function $a(\tau)$. The Jacobian of this system at the points $(\tilde{\tau}, \tilde{a})$ equals $\det D \neq 0$. Therefore, by the Implicit Function Theorem, in a vicinity of $\tilde{\tau}$ there exists a unique continuous vector function $a(\tau)$ such that $a(\tilde{\tau}) = \tilde{a}$. This function is continuously differentiable and

$$\left. \frac{\partial a(\tau)}{\partial \tau} \right|_{\tau = \tau^*} = -D^{-1}B.$$

An immediate calculation now gives

$$\left(\left. \frac{\partial^2}{\partial \tau_j \partial \tau_i} q(\tau, a(\tau), z) \right|_{\tau = \tau^*} \right)_{i,j=1}^{m-1} = E - B^T D^{-1} B.$$

For any fixed $z \in Z$ we have

$$\varphi(\tau, z) = \min_{a \in \mathfrak{A}} q(\tau, a, z) = q(\tau, a(\tau), z)$$

with τ from a vicinity of $\tilde{\tau} = \tilde{\tau}(z)$.

Differentiating this equality twice by τ we receive

$$J(\tilde{\tau}(z), z) = J(\tilde{\tau}, z) = E - B^T D^{-1} B.$$

$\square$

Proof of Theorem 5.3. Let $\tau(z)$ be an arbitrary $K - 1$ times continuously differentiable vector function in a vicinity of a point $z_{(0)}$, $z_{(0)} \in \mathbb{R}^r$, $\tau(z) = (\tau_1(z), \ldots, \tau_{m-1}(z))$. Consider the following auxiliary result.

Lemma 6.1 *Under the condition (b) and with $p = 0$, $l = 0$ the following equalities are valid:*

$$\frac{\partial^t}{\partial z_1^{s_1} \ldots \partial z_k^{s_k}} \left[g\left(\tau_{<I>}(z), z\right) - g\left(\tau(z), z\right) \right] |_{z=z_{(0)}} = 0,$$

for $k \geq 1$, $s \in S_t$, where $I = I_t$, $t = 1, 2, \ldots, K - 1$.

Proof of Lemma 6.1. At first, consider $k = 1$.

Since

$$\frac{\partial}{\partial z} g\left(\tau(z), z\right) = \frac{\partial}{\partial \tau} g(\tau, z)|_{\tau=\tau(z)} \times \tau'(z) + \frac{\partial}{\partial z} g(\tau, z)|_{\tau=\tau(z)}$$

we receive for $t = 1, \ldots, K - 1$

$$
\begin{aligned}
\frac{\partial^t}{\partial z^t} g\left(\tau(z), z\right)|_{z=z_{(0)}} = \\
= t! J_{(0)} \tau_{(t)} + \frac{\partial^t}{\partial z^t} g(\tau_{(0)}, z_{(0)}) + \ldots + \\
+ \sum_{i_1, \ldots, i_t = 1}^{m} \frac{\partial^t}{\partial \tau_{i_1} \ldots \partial \tau_{i_t}} g(\tau_{(0)}, z_{(0)}) \tau_{i_1(1)} \ldots \tau_{i_t(1)} i_1! \ldots i_t!,
\end{aligned}
\tag{17}
$$

where the right-hand side depends only on $\tau_{(0)}, \ldots, \tau_{(t)}$ and does not depend on $\tau_{(t+1)}, \ldots$. Therefore,

$$\frac{\partial^t}{\partial z^t} g\left(\tau(z), z\right)|_{z=z_{(0)}} = \frac{\partial^t}{\partial z^t} g\left(\tau_{(t)}(z), z\right)|_{z=z_{(0)}}.$$

In case $k > 1$, the lemma proof is similar. $\square$

Return to the proof of **Theorem 5.3**. Let $k = 1$, $l = 0$. Note that in the right-hand side of equality (17) only the first term depends on $\tau_{(t)}$ as the other ones depend only on $\tau_{(s)}$, $s \leq t - 1$. Since $g(\tau^*(z), z) \equiv 0$ in a vicinity of $z_{(0)}$, then

$$-\frac{\partial^t}{\partial z^t} g\left(\tau^*_{<t-1>}(z), z\right)|_{z=z_0} = t! J_{(0)} \tau^*_{(t)}.$$

For $k > 1$, $l \neq 0$, the proof is similar. $\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Proof of Theorem 5.4. At first, consider $l = 0$. Note that

$$(g(\tau_{<I>}(z), z))_{(s+q)} = \sum_{w+v=s+q} a_{(w)} \tilde{g}(\tau_{<I>}(z), z)_{(v)} \qquad (18)$$

for any collection of indexes I,

$$\tau_{<I>}(z) = \sum_{s \in I} \tau_{(s)}(z - z_{(0)})^s.$$

For $w = q$, vector s is the only vector v, such that $w + v = s + q$. Let $s \in \hat{S}_n$, $I = \hat{I}_n$. Note that for $w \neq q$ any vector v, such that $w + v = s + q$, belongs to set $\hat{S}_t$, $t \leq n - 1$. Whence it follows that the right-hand side of equality (18) has the form

$$a_{(q)} \tilde{g} \left(\tau^*_{<\hat{I}_n>}(z), z \right)_{(s)}.$$

It can be verified by direct calculation that $J_{(q)} = a_{(q)} \tilde{J}_{(0)}$. Therefore, Proposition 5.4 is valid at $l = 0$. For arbitrary l, its validity can be verified by direct calculation. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Proof of Theorem 3.1.

Consider a vector function $\tilde{\tau}(z) = (\tilde{\tau}_1(z), \ldots, \tilde{\tau}_{m-1}(z))^T$, $\tilde{\tau}(z) : Z \to R^{m-1}$ such that $\xi_{\tilde{\tau}}$ with $\tilde{\tau} = \tilde{\tau}(z)$ is a saturated locally D-optimal design at the point $\Theta^{0T} = \left(\Theta_1^{0T}, (\Psi^{-1}(z))^T \right)$. This function should satisfy equation (3) and due to the Implicit Function Theorem (Gunning, Rossi, 1965) we need only to prove that the Jacobi matrix is invertible.

Suppose, oppositely that it is not the case. Then there exists a vector $d \in R^{m-1}$, $d \neq 0$ such that $d^T B = 0$ and therefore

$$\sum_{s=1}^{m-1} \left[f_i(x_j^*) f_j'(x_s^*) + f_i'(x_s^*) f_j(x_s^*) \right] d_s = 0, \qquad (19)$$

$i, j = 1, \ldots, m$, $(i, j) \neq (m, m)$, $x_s^* = \bar{\tau}_s(z)$, $f_i(x) := f_i(x, z)$, $i = 1, \ldots, m$, $s = 1, \ldots, m - 1$.

Note the equation (19) holds also for $(i, j) = (m, m)$. Really, since

$$\xi_{\bar{\tau}} = \begin{pmatrix} x_1^* & \cdots & x_{m-1}^* & b \\ 1/m & \cdots & 1/m & 1/m \end{pmatrix}$$

23

is a saturated locally D-optimal design, we have

$$\frac{\partial}{\partial x_s} \det M(\xi_{\bar{\tau}}, z) = \sum_{i,j=1}^{m} (f_i(x_s) f_j(x_s))' d_{ij} = 0 \tag{20}$$

where $d_{ij} = (M^{-1}(\xi_{\bar{\tau}}, z))_{i,j}$, $s = 1, \ldots, m-1$.

Multiplying (20) by d_s and summing the results we receive

$$\sum_{i,j=1}^{m} \left(\sum_{s=1}^{m-1} (f_i(x_s) f_j(x_s))' d_s \right) d_{ij} = 0.$$

Substituting (19) tp these equations we receive

$$\left(\sum_{s=1}^{m-1} \left(f_m^2(x_s) \right)' d_s \right) d_{mm} = 0.$$

Since $(M(\xi_{\bar{\tau}}, z))^{-1}$ is a negative definite matrix

$$d_{mm} = e_m^T (M(\xi_{\bar{\tau}}, z))^{-1} e_m \neq 0, \ e_, = (0, \ldots, 0, 1)^T$$

and thus (19) holds for $(i, j) = (m, m)$.

Define a vector q by the following equality

$$q^T f(x) = \det \begin{pmatrix} f_1(x_1^*) & \ldots & f_m(x_m^*) \\ \cdots \cdots \cdots \\ f_1(x_{m-1}^*) & \ldots & f_m(x_{m-1}^*) \\ f_1(x) & \ldots & f_m(x) \end{pmatrix}.$$

Certainly $q^T(x_i^*) = 0$, $i = 1, \ldots, m-1$ and we receive from (19) that

$$\sum_{s=1}^{m-1} q^T f'(x_s^*) f_j(x_s^*) d_s = 0, \ j = 1, \ldots, m.$$

Due to Assumption **A1** we have $q^T f'(x_s^*) \neq 0$, $s = 1, \ldots, m$. And therefore

$$L_{(t)} \alpha = 0, \ t = 1, \ldots, m, \tag{21}$$

where $\alpha = \left(d_s q^T f'(x_s^*) \right)_{s=1}^{m-1}$, $L_{(t)}$ is received from the matrix $\left(f_i(x_j^*) \right)_{i,j=1}^{m, m-1}$ by rejecting t-th line. It follows from (21) that $\det L_{(t)} = 0$, $t = 1, \ldots, m$ and it implies $\det \left(f_i(x_j^*) \right)_{i,j=1}^{m} = 0$. But the last equality is impossible. $\square$

Proof of Propositions 4.2 and 4.3. The existence of a locally optimal design is evident.

Let z be fixed, $\xi^* = \{x_1^*, \ldots, x_n^*; \mu_1^*, \ldots, \mu_n^*\}$ be a locally optimal design, $0 \leq x_1^* < \ldots < x_n^* \leq d$.

Denote
$$g(x) = f^T(x) M^{-1}(\xi^*, \Theta) f(x),$$

where $f(x) = (f_1(x, \Theta), \ldots, f_m(x, \Theta))^T$ and $\Theta = \Theta^{(0)}$ is fixed.

Due to the Kiefer–Wolfowitz equivalence theorem

$$g(x) \leq m, \ g(x_i^*) = m, \ i = 1, \ldots, n.$$

It follows that
$$g'(x_i^*) = 0, \ i = 2, \ldots, n - 1,$$

while if $x_1^* \neq 0$ then $g'(x_1^*) = 0$ and if $x_n^* \neq 0$ then $g'(x_n^*) = 0$. The function $\tilde{g}(x) = g(x) - m$ is of the form

$$\tilde{g}(x) = P(x)/Q^4(x),$$

where $Q(x) = \prod_{i=1}^{k}(x + \theta_{i+k})$, $P(x)$ is a polynomial of degree $2m$. By analyzing this expression we can verify that $x_1^* = 0$, $n = m$, $x_m^* < d$ for sufficiently large d and all elements of the matrix E are negative. Thus due to the remark after Theorem 5.2 the condition (a) is fulfilled for any $z_{(0)} \in Z$.

Using the formula for the Vandermonde determinant and elementary operations under columns of the matrix $(f_i(x_j))_{i,j=1}^m$ it can be checked that

$$\varphi(\tau, z) = [\psi(z)]^{2/m} \left(\prod_{1 \leq j < j \leq m}(x_i - x_j) / \prod_{i=1}^m Q^2(x) \right)^{2/m}.$$

Now it is evident that assumptions **A1–A4** take place. This completes the proof of Proposition 4.3.

In order to prove Proposition 4.4 let

$$q(x) = \prod_{i=2}^m (x - x_i^*) = \sum_{j=0}^{2k-1} q_j x^{2k-1-j}, \ q_0 = 1.$$

Since
$$\frac{1}{2} \frac{q''(x_i^*)}{q'(x_i^*)} = \sum_{j \neq i} \frac{1}{x_i^* - x_j^*}$$

25

(Fedorov, 1972, Ch. 2.3) the equation $\tilde{g}(\tau, z) = 0$ attains the form

$$q^{''}(x)x\prod_{l=1}^{k}(x + \lambda_l) + 2q^{'}(x)\left(\prod_{l=1}^{k}(x + \lambda_l) - 2x\sum_{i=1}^{k}\prod_{l\neq i}(x + \lambda_l)\right) = 0.$$

Passing to the limit with $z \to z_{(0)}$ we receive

$$\begin{aligned} q^{''}(x)x(x + 1) + 2q^{'}(x)(x(1 - 2k) + 1) &= \kappa q(x), \\ \kappa = (m - 1)(m - 2) + 2(m - 1)(1 - 2k). \end{aligned} \tag{22}$$

Equating coefficients at equal degrees in the right and the left hand sides of this equality we see that coefficients of $q(x)$ can be calculated by recurrent formulas in a unique way.

Acknowledgments

References

[1] Bezeau, M. and Endrenyi, L. (1986). Design of experiments for the precise estimation of dose-response parameters: the Hill equation. *Journal of Theoretical Biology*, 123, 4, 415–430.

[2] Dette, H., Melas, V.B. (2002). E-optimal designs in Fourier regression models on a partial circle. *Mathematical Methods of Statistics*, 11, N 3, pp. 259-296.

[3] Dette, H., Melas, V.B. (2003). Optimal designs for estimating individual coefficients in Fourier regression models. *Ann. Stat.*, 31, N 5, 1669–1692.

[4] Dette, H., Melas, V.B.,Pepelyshev, A. (2002). D-optimal designs for trigonometrical regression models on a partial circle. *Annals of the Institute of Statistical Mathematics*, 54, No. 4,945–959.

[5] Dette, H., Melas, V.B., Pepelyshev A. (2004). Optimal designs for estimating individual coefficients − a functional approach . *Journal of Statistical Planning and Inference*, 118, 201–209.

[6] Dette, H., Melas, V.B., Wong, W.K. (2003). Locally D-optimal designs for exponential regression. Preprint Ruhr-Universität Bochum. http://www.ruhr-uni-bochum.de/mathematik3/preprint.htm

[7] Dette, H., Melas, V.B., Wong, W.K. (2004). Optimal design for goodness-of-fit of the Michaelis–Menten enzyme kinetic function. Preprint Ruhr-Universität Bochum. http://www.ruhr-uni-bochum.de/mathematik3/preprint.htm

[8] Jennrich, R.J. (1969). Asymptotic properties of non-linear least squares estimators. *The Annals of Mathematical Statistics*, 40, 633–643.

[9] Fedorov, V.V. (1972). *Theory of Optimal Experiments*. Academic Press, New York.

[10] Gunning, R.C., Rossi, H. (1965). *Analytic Functions of Several Complex Variables*. Prentice Hall, New York.

[11] Han C., Chaloner K. (2003). *D*- and *C*-optimal Designs for Exponential Regression Models Used in Viral Dynamics and Other Applications. *Journal of Statistical Planning and Inference*, 115, 585–601.

[12] He, Z., Studden, W.J., Sun, D. (1996). Optimal designs for rational models. *Ann. Statist.* 24, 2128–2147.

[13] Karlin, S. Studden W. (1966). *Tchebysheff Systems: with Application in Analysis and Statistics*. New York, Wiley.

[14] Kiefer, J. (1985). *Collected Papers*. Springer, New York.

[15] Melas V.B. (1978). Optimal designs for exponential regression. *Mathematische Operationsforshung und Statistik*, Ser. Statistics, 9, 45–59.

[16] Melas V.B.(2000). Analytic theory of *E*-optimal designs for polynomial regression. In: Advances in Stochastic Simulation Methods (eds. N. Balakrishnan, V.B. Melas, S. Ermakov), Birkhäuser, Boston, 85-115.

[17] Melas V.B. (2001). Analytical properties of locally *D*-optimal designs for rational models. In: Atkinson A.C., Hackel P., Müller W.J. (eds.) *MODA 6 — Advances in Model-Oriented Design and Analysis*, Physica-Verlag, Heidelberg, 201–210.

[18] Müller, Ch.H. (1995). Maximin efficient designs for estimating nonlinear aspects in linear models. *Journal of Statistical Planning and Inference*, 44, 117–132.

[19] Petrushev, P.P., Popov, V.A. (1987). *Rational Approximations of Real Functuons*. Cambridge Univ.Press.

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