Chapter 2

The Functional Approach

This chapter is devoted to studying optimal designs for a wide class of nonlinear regression models on the basis of a functional approach. This class includes exponential and rational models as well as many particular models of the Chebyshev type used in microbiology and other fields of experimental research.

We consider designs that are locally D-optimal or maximin efficient D-optimal among designs with the number of points equal to the number of parameters. In many cases, such designs prove to be optimal or maximin efficient among all approximate designs.

Support points of such designs are considered here as implicit functions on the initial value of the nonlinear parameters or on characteristics of sets containing, by the assumption, the true parameter value. A corresponding equation system is derived and is called the basic equation system or the basic (vector) equation. Studying this system allows one to prove that the functions are real analytic and therefore can be represented by a Taylor series under natural conditions. Recurrent formulas for computer-calculating the Taylor coefficients are introduced.

2.1 Introduction

Most results in the modern regression design theory were obtained for linear models with a fixed design region (see Fedorov, 1972; Silvey, 1980; Kiefer, 1985; Pukelsheim, 1993). However, many models of practical importance are nonlinear models (see, e.g., Seber and Wild, 1989). The commonly used approach for experimental design in such models consists of their linearization in a vicinity of some initial values of the nonlinear parameters and application of locally optimal designs, briefly discussed in the previous chapter. In spite of such designs are usually depending on the initial values, they can be used if a reliable knowledge about the parameters is available. These designs are also used in studying more complicated approaches: maximin and Bayesian ones (see Section 1.7).

Even if linear models are implemented, the design region often cannot be considered as fixed. For example, in many microbiological studies (see Pirt, 1984; Dette, Melas, and Strigul, 2005), the design region is a time interval and can be chosen by an experimentator in different ways. The introduction of design intervals with variable bounds can be considered also as an artificial method for investigations of the structure of optimal designs.

In the present chapter, we will consider nonlinear models given at a design interval. Our basic method here is the functional approach introduced in Melas (1978) for studying exponential nonlinear models and our aim is to apply it to a wider class of models.

The main idea of this approach consists of studying optimal design points and weights as implicitly given functions of the bound of the design interval and/or nonlinear parameters of the model. These functions can be investigated on the basis of the Implicit Functional Theorem formulated in Section 1.8 (see also Gunning and Rossi (1965)). In particular, in many cases these functions prove to be real analytic which enables one to present them by segments of the Taylor series. We will introduce here general recurrent formulas for constructing such series and discuss their applications for studying properties of optimal designs.

The functional approach seems to be useful when an explicit analytical form of optimal designs is not available. It can be considered as an alternative or useful addition to merely numerical methods. It is worth mentioning that similar approaches are well known in many fields of mathematics and its application. For example, representing indefinite integrals by a power series is the recognized technique of their calculation, and coefficients of such series are tabulated and given in textbooks. However, in the field of experimental design the functional approach is relatively new. References to existing literature will be given throughout the book.

In Section 2.2*, we will introduce the basic ideas of the functional approach using exponential models (nonlinear by parameters) as an example. Section 2.3 contains a list of assumptions justifying the implementation of the functional approach and formulates without proofs the main theoretical results. Section 2.4 is devoted to studying the basic equation. It is also introduces general recurrent formulas for calculating the Taylor coefficients. The application of the theory to the three-parameter logistic model is given in Section 2.5. All lengthy proofs are deferred to Section 2.6.

*Note that in Section 2.2 and in Sections 2.3–2.6 a part of materials are taken from Melas, V.B. (2005). On the functional approach to optimal designs for nonlinear models. *J. Statist. Plan. and Inference*, **132**, 93–116.

2.2 Basic Ideas of the Functional Approach

In this section we will introduce some basic ideas of the approach. We will consider regression models given by linear combinations of unknown exponentials as a typical example of nonlinear models.

In order to make the explanation more apparent, all technically difficult mathematical results will be only formulated and their proofs will be given in further sections.

Let us restrict our attention by the D-criterion and study locally D-optimal designs and maximin efficient D-optimal designs.

As it was discussed in Section 1.7, the first of the problems has some independent interest. It is also a necessary step for investigating the second problem.

2.2.1 Exponential regression models

Let us consider the models given by relations

$$Y_j = \sum_{i=1}^{k} a_i e^{-\lambda_i x_j} + \varepsilon_j,\ j = 1, \ldots, N, \tag{2.1}$$

where $Y_1, \ldots, Y_N$ are experimental results, $a_1, \ldots, a_k$ and $\lambda_1, \ldots, \lambda_k$ are the parameters to be estimated; and

$$a_i \neq 0,\ \lambda_i > 0,\ i = 1, \ldots, k,\ \lambda_i \neq \lambda_j\ (i \neq j), \tag{2.2}$$

$\varepsilon_1, \ldots, \varepsilon_N$ are independent and identically distributed random values (experimental errors) with zero mean ($E\varepsilon_i = 0$) and the variance $E\varepsilon_i^2 = \sigma^2 > 0$, and $x_1, \ldots, x_N \in [0, \infty)$ are observation points.

Let us assume that k is known and the problem consists of an optimal choice of observation points in order to estimate the parameters as accurately as possible for a given number of possible observations at the interval $[0, \infty)$.

The model (2.1) is of a great theoretical and practical interest. It is often used in chemical and biological investigations (see, e.g., Becka and Urfer (1996) and Han and Chaloner (2003)).

A discrete probability measure

$$\xi = \begin{pmatrix} x_1 & \ldots & x_n \\ \omega_1 & \ldots & \omega_n \end{pmatrix}, \tag{2.3}$$

where $0 < x_1 < \cdots < x_n$ are support points and $\omega_i > 0$, $i = 1, \ldots, n$, and $\sum \omega_i = 1$ are weight coefficients, will be called the (approximate) experimental design.

Let we have an opportunity to realize N experiments. We will say that the experiments are performed in accordance with the design (2.3) if r_i

observations are performed in points x_i $(i = 1, \dots, n)$, where

$$r_i = \lfloor \omega_i N \rfloor \text{ or } \lfloor \omega_i N \rfloor + 1,$$

and $\lfloor a \rfloor$ denotes the integer part of a, in such a way that $\sum r_i = N$.

Set $\theta = (a_1, \lambda_1, \dots, a_k, \lambda_k)^T$. Denote by $\hat{\theta}(N)$ the least squares estimator for θ obtained from the results of N experiments in accordance with a design of the form (2.3); that is,

$$\hat{\theta} = \hat{\theta}(N) = \arg \min_{\theta \in R^{2k}} \sum_{i=1}^{n} \sum_{j=1}^{r_i} \left[Y_{ij} - \eta(x_i, \theta)\right]^2,$$

where

$$\eta(x, \theta) = \sum_{i=1}^{k} a_i e^{-\lambda_i x}$$

and Y_{ij} is the result of the j-th experiment in the point x_i.

Let θ^* denote the true value of θ in the model (2.1). It can be shown by verification of regularity conditions of the Jennrich theorem (Jenrich, 1969) that with $n \geq 3$ and $N \to \infty$, the covariance matrix of the vector $(\hat{\theta}(N) - \theta^*)/\sqrt{N}$ tends to the matrix

$$\sigma^2 \left[\int f(x, \theta) f^T(x, \theta) \xi(dx)\right]^{-1}, \tag{2.4}$$

where

$$\begin{aligned} f(x, \theta) &= \frac{\partial}{\partial \theta} \eta(x, \theta), \\ \int g(x) \xi(dx) &= \sum_{i=1}^{n} g(x_i) \omega_i, \\ \theta &= \theta^*. \end{aligned}$$

The matrix

$$\left(\sum_{s=1}^{n} \frac{\partial \eta(x_s, \theta)}{\partial \theta_i} \frac{\partial \eta(x_s, \theta)}{\partial \theta_j} \omega_s \right)_{i,j=1}^{2k}$$

is usually called the Fisher information matrix.

By immediate application of Binet–Cauchy's formula to the determinant of this matrix, we obtain

$$\begin{aligned} &\det\left(\textstyle\int f(x, \theta) f^T(x, \theta) \xi(dx)\right) \\ &= a_1^2 \dots a_k^2 \textstyle\sum_{1 \leq i < \dots < i_{2k} \leq n} \left(\prod_{s=1}^{2k} \omega_{i_s}\right) \det^2 \left(\psi_l(x_{i_j})\right)_{l,j=1}^{2k}, \end{aligned} \tag{2.5}$$

where

$$\psi_1(x) = \frac{\partial}{\partial a_1}\eta(x,\theta) = e^{-\lambda_1 x},$$
$$\psi_2(x) = \frac{\partial}{\partial \lambda_1}\eta(x,\theta) = -xe^{-\lambda_1 x},$$
$$\vdots$$
$$\psi_{2k-1}(x) = \frac{\partial}{\partial a_k}\eta(x,\theta) = e^{-\lambda_k x},$$
$$\psi_{2k}(x) = \frac{\partial}{\partial \lambda_k}\eta(x,\theta) = -xe^{-\lambda_k x}.$$

Let us restrict ourselves by the D-criterion of optimality (for other criteria, we can proceed in a similar way). A design is called *D-optimal* if it maximizes the determinant of the information matrix. The problem is to find a design maximizing the determinant among all possible (approximate) designs. Note that with $a_i \neq 0$, $i = 1, \dots, k$, values of $a_1, \dots, a_k$ do not influence the solution of this problem (since they involve only in the multipliers $a_1^2, \dots, a_k^2$). Therefore, we can assume in the following that $a_1^2 = \dots = a_k^2 = 1$.

However, the design maximizing the value (2.5) depends, generally speaking, on the value $\Lambda = (\lambda_1, \dots, \lambda_k) = \Lambda^* = (\lambda_1^*, \dots, \lambda_k^*)$. Such a dependence is the main feature of all nonlinear models.

There are several ways to overcome this difficulty. Let us begin with the locally optimal approach (introduced by Chernoff (1953)). This approach consists of the replacement of the unknown value Λ^* by a known approximation for it (an initial guess).

A design will be called locally D-optimal if it maximizes the determinant (2.5) with $a_1 = \dots = a_k = 1$ and $\Lambda = (\lambda_1, \dots, \lambda_k) = \Lambda^{(0)} = (\lambda_1^{(0)}, \dots, \lambda_k^{(0)})$.

Let us set

$$M(\xi, \Lambda) = \int f(x,\theta) f(x,\theta)^T \xi(dx),$$

where $\theta = (1, \lambda_1, \dots, 1, \lambda_k)$ and $\Lambda = (\lambda_1, \dots, \lambda_k)^T$.

Note that this matrix coincides with the information matrix for the corresponding linear model

$$Y = \beta_1 e^{-\lambda_1 x} + \beta_2 x e^{-\lambda_1 x} + \dots + \beta_{2k-1} e^{-\lambda_k x} + \beta_{2k} e^{-\lambda_k x} + \varepsilon, \qquad (2.6)$$

where $\beta_1, \dots, \beta_{2k}$ are parameters to be estimated and $\lambda_1, \dots, \lambda_k$ are assumed to be known.

Now, we should make a very important remark. Note that we assumed $\lambda_i \neq \lambda_j$ $i \neq j$, when we formulated our model. In fact, if $\lambda_i = \lambda_j$ for some $i \neq j$, then the model (2.1) contains no more than $k-1$ terms of the form

$$a_i e^{-\lambda_i x}.$$

However, we restrict ourselves by the models with k terms.

It should be noted that if $\lambda_i = \lambda_j$ for some i and j $(i \neq j)$, then for each of the determinants in the right-hand side of (2.5) the two corresponding columns coincide. Thus, in this case, $\det M(\xi, \Lambda) = 0$ for any design ξ.

However, from a mathematical point of view, it is useful to admit that the value

$$\min_{i \neq j} \left| \lambda_i - \lambda_j \right|$$

can be as small as we like. Moreover, it can be verified [see Melas (1978)] that the function

$$V(\xi, \Lambda) = (\det M(\xi, \Lambda)) \Big/ \prod_{i<j} (\lambda_i - \lambda_j)^8 \tag{2.7}$$

can be codefined with preserving continuity at the set of all positive values $\lambda_1, \ldots, \lambda_k$.

Now, we are prepared to introduce a more convenient definition.

Definition 2.2.1 A design, maximizing the value (2.7) among all (approximate) designs for an arbitrary fixed vector Λ with positive coordinates will be called *a locally D-optimal design*.

For arbitrary Λ such that $\lambda_i \neq \lambda_j$ $(i \neq j)$, this definition corresponds to the usual definition of locally D-optimal designs.

Note that designs that maximize the limit of (2.7) with $\Lambda \to \Lambda_\gamma = \gamma(1, \ldots, 1)$ that is locally D-optimal designs for points $\Lambda = \Lambda_\gamma$ will play an important role in the following consideration. Due to the continuity arguments these designs will be nearly optimal for all vectors Λ whose coordinates are close enough to each other.

We will construct and study locally D-optimal designs in the next subsection.

It should be noted that locally D-optimal (LD)designs depend on the initial vector $\Lambda = \Lambda^{(0)}$ and could be not very efficient if this vector is far from the vector of true parameter values Λ^*. However, the design could be implemented in a sequential manner. One can take $\Lambda = \Lambda^{(0)}$, construct an LD design for this vector, and realize N_1 experiments in accordance with this design. Then one can construct the LS (least squares) estimator $\hat{\theta} = \hat{\theta}(N_1)$ and take the parameter vector $\hat{\Lambda}^{(1)} = \hat{\Lambda}(N_1)$ in order to construct the new LD design. By repeating this procedure several times, we will obtain a design close to the LD design with $\Lambda = \Lambda^*$.

The described procedure (see, e.g., Silvey (1980) for more accurate explanation) cannot be appropriate if we need to have a design for all experiments in advance. An alternative to such a sequential implementation of LD design consists of using a minimax approach (see Section 1.7 for a more detailed discussion).

Let us consider a reasonable version of the minimax approach.

Assume that for the vector Λ^*, a set Ω of its possible values is given. In particular, such a set can be obtained from preliminary experiments or by theoretical consideration of the underlying real problem. From a practical point of view, the following type of set seems to be of a great interest:

$$\Omega = \Omega(\delta) = \{\Lambda; (1-\delta)x_i \leq \lambda_i \leq (1+\delta)c_i,\, i = 1, \ldots, k\}\,, \tag{2.8}$$

where c_i is an approximation to λ_i^*, $i = 1, \ldots, k$, and the value $\delta \in (0,1)$ can be interpreted as a relative error of this approximation.

Note that the intervals $[(1-\delta)c_i, (1+\delta)c_i]$ can be overlapped and even can coincide with each other.

From a methodical point of view, it is very convenient that the set (2.8) under fixed $c_1, \ldots, c_k$ is determined by a single parameter δ.

Let us call a design a maximin efficient D-optimal design if it maximizes the value

$$\min_{\Lambda\in\Omega} \left[\frac{V(\xi,\Lambda)}{V(\xi(\Lambda),\Lambda)}\right]^{1/m},\ m = 2k, \tag{2.9}$$

where $\xi(\Lambda)$ is a LD design, $\Omega = \Omega(\delta)$ is determined by (2.8).

Note that the minimum here is achieved at some values $\bar\Lambda \in \Omega$ since Ω is a bounded and closed set.

The value (2.9) for a given design will be called the minimal efficiency.

Note that

$$\left[\frac{V\xi,\Lambda)}{V(\xi(\Lambda),\Lambda)}\right]^{1/m} = \left[\frac{\det M(\xi,\Lambda)}{\det M(\xi(\Lambda),\Lambda)}\right]^{1/m}$$

if Λ satisfies the restriction $\lambda_i \neq \lambda_j$ $(i \neq j)$.

If we perform N experiments in accordance with a design ξ, then the volume of a confidence ellipsoid for LS estimates will be proportional to

$$\left(\frac{1}{\sqrt{N}}\right)^m \sqrt{\det M(\xi,\Lambda)}$$

(see, e.g., Pukelsheim (1993)).

Thus, the minimal efficiency of a given design is equal to the ratio N/N^*, where N is the number of experiments along the design ξ needed for obtaining estimates with a given accuracy and N^* is the similar number for a LD design.

In the following subsections we will demonstrate opportunities of the functional approach to constructing and studying LD and maximin efficient D-optimal designs.

2.2.2 Locally D-optimal designs

It is easy to check that if the number of support points of a design ξ is less than the number of parameters to be estimated $(n < 2k)$, then $\det M(\xi,\Lambda) = 0$. By this reason, the designs with $n = 2k$ is usually called designs with minimal support. In the following we restrict our attention by such designs, and in Chapter 6, it will be shown that LD designs for the exponential model (2.1) usually belong to this class of designs. Designs that are LD in the class of designs with minimal support will be called, for brevity, LDMS designs.

An immediate calculation shows that with $n = 2k$,

$$M(\xi, \Lambda) = F^T W F,$$

where $W = \text{diag}\{\omega_1, \ldots, \omega_{2k}\}$, $F = (\psi_l(x_j))_{l,j=1}^{2k}$, and $\psi_l(x)$ are defined in (2.8). Therefore,

$$\begin{aligned}\det M(\xi, \Lambda) &= \textstyle\prod_{i=1}^{2k} \omega_i \det^2 F \\ &\le \left(\frac{\sum \omega_i}{2k}\right)^{2k} \det^2 F = \left(\left(\tfrac{1}{2k}\right)^{2k}\right) \det^2 F,\end{aligned}$$

whereas the equality takes place if and only if $\omega_i = \frac{1}{2k}$, $i = 1, \ldots, 2k$. Thus LDMS designs have the form

$$\xi = \begin{pmatrix} x_1 & \ldots & x_m \\ \frac{1}{m} & \ldots & \frac{1}{m} \end{pmatrix}, \; 0 \le x_1 < \cdots < x_m, \; m = 2k,$$

that is, all weight coefficients in such designs are the same.

Let us prove that in each of LDMS designs $x_1 = 0$. Set

$$\xi_\Delta = \begin{pmatrix} x_1 + \Delta & \ldots & x_m + \Delta \\ \frac{1}{m} & \ldots & \frac{1}{m} \end{pmatrix}, \; F_\Delta = (\psi_l(x_j + \Delta))_{l,j=1}^m.$$

Consider the determinant

$$\det F_\Delta = \det \begin{pmatrix} e^{-\lambda_1(x_1+\Delta)} \ldots e^{-\lambda_1(x_m+\Delta)} \\ -(x_1+\Delta)e^{-\lambda_1(x_1+\Delta)} \ldots -(x_m+\Delta)e^{-\lambda_1(x_m+\Delta)} \\ e^{-\lambda_k(x_1+\Delta)} \ldots e^{-\lambda_k(x_m+\Delta)} \\ -(x_1+\Delta)e^{-\lambda_k(x_1+\Delta)} \ldots -(x_m+\Delta)e^{-\lambda_k(x_m+\Delta)} \end{pmatrix}.$$

Let us add the first line multiplied by Δ to the second line, ..., and the $(2k-1)$-st line multiplied by Δ to the $(2k)$-th line. Then let us extract from each of the lines the multiplies of the form $e^{-\lambda_i \Delta}$, $i = 1, \ldots, k$. In this way, we obtain

$$\det F_\Delta = e^{-2(\sum_{i=1}^k \lambda_i)\Delta} \det F,$$

and with $\Delta < 0$,

$$\det{}^2 F_\Delta > \det{}^2 F.$$

Thus, with $x_1 > 0$, a design ξ cannot be LDMS since with $\Delta = -x_1$,

$$\det M(\xi_\Delta, \Lambda) = \left(\frac{1}{m}\right)^m \det{}^2 F_\Delta > \det M(\xi, \Lambda).$$

Therefore, for any LDMS design, we have $x_1 = 0$.

Let us introduce the following notation:

$$\tau = (\tau_1, \dots, \tau_{m-1}) = (x_2, \dots, x_m),$$

$$\xi_\tau = \begin{pmatrix} 0 & x_2 & \dots & x_m \\ \frac{1}{m} & \frac{1}{m} & \dots & \frac{1}{m} \end{pmatrix},$$

$$\varphi(\tau, \Lambda) = (V(\xi_\tau, \Lambda))^{1/m},$$

$$R_+^s = \{u \,:\, u \in R^s, u = (u_1, \dots, u_s);\; u_i > 0,\; i = 1, \dots, s\}.$$

Note that there exists a one-to-one correspondence between vectors $\tau \in R_+^{m-1}$ and designs of the form

$$\xi = \xi_\tau = \begin{pmatrix} 0 & x_2 & \dots & x_m \\ \frac{1}{m} & \frac{1}{m} & \dots & \frac{1}{m} \end{pmatrix}.$$

The problem of LDMS designs is now reduced to the maximization of the function $\varphi(\tau, \Lambda)$ by $\tau \in R_+^{m-1}$ under a fixed Λ, where $\Lambda = (\lambda_1, \dots, \lambda_k)$, $\lambda_i > 0$, and $i = 1, \dots, 1$.

Since

$$\varphi(\tau, \Lambda) = C(\Lambda)(\det F)^{2/m},$$

$F = (\psi_l(x_j))_{l,j=1}^{2m}$, where $C(\Lambda)$ does not depend on τ and each of elements of F tends to zero with $x_m \to \infty$, then the maximum of $\varphi(\tau, \Lambda)$ by $\tau \in R_+^{m-1}$ is achieved in an inner point of R_+^{m-1} for which $0 < \tau_1 < \dots < \tau_m$. Due to the known necessary conditions for extremum points in order for a design $\xi = \xi_{\tau*}$ to be an LDMS design, it is necessary that with $\tau = \tau*$, the following equalities be satisfied,

$$\frac{\partial}{\partial \tau_i} \varphi(\tau, \Lambda) = 0,\; i = 1, \dots, m-1. \tag{2.10}$$

Consider the case $k = 1$. In this case,

$$\det M(\xi_\tau, \Lambda) = \left[\frac{1}{2} \det \begin{pmatrix} 1 & e^{-\lambda_1 x_2} \\ 0 & -x_2 e^{-\lambda_1 x_2} \end{pmatrix}\right]^2 = \frac{1}{4} x_2^2 e^{-2\lambda_1 x_2},$$

$$\varphi(\tau, \Lambda) = [\det M(\xi_\tau, \Lambda)]^{1/2} = \frac{1}{2} x_2 e^{-\lambda_1 x_2} = \frac{1}{2} \tau_1 e^{-\lambda_1 \tau_1}.$$

Equalities (2.10) assume the form of the single equation

$$\frac{\partial}{\partial \tau_1} (\tau_1 e^{-\lambda_1 \tau_1}) = e^{-\lambda_1 \tau_1}(1 - \lambda_1 \tau_1) = 0.$$

The unique solution of this equation under fixed λ_1 is

$$\tau_1 = 1/\lambda_1.$$

Thus, in the case $k = 1$, there exists the unique LDMS design

$$\xi* = \xi_{\tau*} = \begin{pmatrix} 0 & 1/\lambda \\ 1/2 & 1/2 \end{pmatrix}.$$

It can be proved (see Chapter 6) that this design is a LD design among all approximate designs.

In the case $k > 1$, it seems impossible to find such an explicit solution of the problem for arbitrary vectors Λ. However, we can find an explicit solution for points Λ of the form $\Lambda = (\gamma, \ldots, \gamma)$, where $\gamma > 0$ is an arbitrary given number.

In fact, in this case,

$$V(\xi_\tau, \Lambda_\gamma) = \lim_{\Lambda \to \Lambda_\gamma} \det M(\xi_\tau, \Lambda) / \prod_{i<j} (\lambda_i - \lambda_j)^8.$$

In order to calculate this limit, use the expansion of the exponential into the Taylor series and elementary properties of the determinant. In Melas (1978) it was proved that this limit is equal to

$$\left(\frac{1}{m}\right)^m e^{-\gamma \sum_{i=2}^m x_i} \prod_{i<j} (x_j - x_i)^2. \tag{2.11}$$

It is easy to check that the value (2.11) coincides with the value of the determinant of the information matrix for linear (by parameters) regression model

$$E(Y|x) = e^{-\gamma x} \sum_{i=1}^m \beta_i x^{i-1},$$

where $\gamma > 0$ is a given number and $\beta_1, \ldots, \beta_m$ are the parameters to be estimated.

As it is known (see Karlin and Studden (1966, Chap. X)), (2.11) has the unique extremal point

$$\tau^* = (x_2^*, \ldots, x_m^*) = \frac{1}{\gamma}(\gamma_1, \ldots, \gamma_{m-1}),$$

where $\gamma_1, \ldots, \gamma_{m-1}$ are the roots of the Laugerre's polynomial of degree $m-1$ with the associated parameter 1. Thus, we know the unique solution of the equation system (2.10) under $\Lambda = \Lambda_\gamma$. For the case of an arbitrary Λ, it can be proved (see Melas (1978)) that the equation system (2.10) has a unique solution. Denote this solution by $\tau^* = \tau^*(\Lambda)$. With arbitrary k, the unique LDMS design is

$$\xi* = \xi * (\Lambda) = \xi_{\tau^*(\Lambda)}.$$

Considering the determinant of the matrix F, it is easy to check that for any scalar $h \neq 0$,

$$\varphi(\tau, h\Lambda) = h\varphi\left(\frac{\tau}{h}, \Lambda\right).$$

Therefore, $\tau^*(h\Lambda) = \tau^*(\Lambda)/h$ and we can restrict our attention to vectors Λ with $\sum_{i=1}^{k} \lambda_i = k$. It allows one to reduce the number of parameters.

Let us introduce the new parameters

$$z = (z_1, \ldots, z_{k-1})^T, \; z_i = 1 - \lambda_i, i = 1, \ldots, k-1.$$

Note that, for $k = 2$, the number of new parameters is equal to 1. Note also that with $\sum_{i=1}^{k} \lambda_i = k$, there exists the one-to-one correspondence between the set of new parameters and the set of vectors Λ:

$$\lambda_i = 1 - z_i, \; i = 1, \ldots, k-1; \; \lambda_k = k - \sum_{i=1}^{k-1} \lambda_i = 1 + \sum_{i=1}^{k-1} z_i.$$

Denote

$$\begin{aligned} \bar{\varphi}(\tau, z) &= \varphi(\tau, \Lambda(z)), \\ g_i(\tau, z) &= \frac{\partial}{\partial \tau_i} \bar{\varphi}(\tau, z), \; i = 1, \ldots, m-1, \\ g(\tau, z) &= (g_1(\tau, z), \ldots, g_{m-1}(\tau, z))^T. \end{aligned} \tag{2.12}$$

Now the equation system (2.10) can be written as the vector equation

$$g(\tau, z) = 0. \tag{2.13}$$

This equation determines the vector function

$$z \to \bar{\tau}^*(z) = \tau^*(\Lambda(z))$$

implicitly, which allows to apply the Implicit Function Theorem (see Section 1.8).

We will now present an extended formulation of this theorem for the vector function $g(\tau, \Lambda)$ of a general form (not necessary connected with the design problem considering here).

Let $g(\tau, z)$, $\tau \in R^{m-1}$, $z \in R^{k-1}$, be an arbitrary vector function $g = (g_1, \ldots, g_{m-1})^T$ with the following properties:

(i) $g(\tau, z)$ is a real analytic vector function in the point $(\tau_{(0)}, z_{(0)})$ (this means that the component of this vector function can be expanded into a convergent multivariate Taylor series in the point).

(ii) $g(\tau_{(0)}, z_{(0)}) = 0$.

(iii) The Jacobi matrix

$$J_{(0)} = \left(\frac{\partial g_i(\tau, z)}{\partial \tau_j}\right)_{i,j=1}^{m-1}\Bigg|_{\tau=\tau_{(0)}, z=z_{(0)}}$$

is invertible.

In order to formulate the theorem, let us introduce the following notations. Let $Q(u)$ be an arbitrary (scalar or vector) function of one variable that is infinitely many times differentiable in a point $u_{(0)}$. Denote

$$Q_{(0)} = Q(u_{(0)}),$$

$$Q_{(3)} = \frac{1}{s!}\frac{d^s}{du^s}Q(u)\Bigg|_{u=u_{(0)}}, \; s = 1, 2, \ldots$$

If the function $Q(u)$ is real analytic in a vicinity of the point $u = u_{(0)}$, then

$$Q(u) = Q_{(0)} + \sum_{s=1}^{\infty} Q_{(s)}(u - u_{(0)})^s$$

in this vicinity.

In the multidimensional case $u = (u_1, \ldots, u_{k-1})$, it is necessary to interpret s as the multi-index $s = (s_1, \ldots, s_{k-1})$ and denote

$$Q_{(s)} = \frac{1}{s_1!}\cdots\frac{1}{s_{k-1}!}\frac{\partial^{s-1}}{\partial u_1^{s_1}}\cdots\frac{\partial^{s_{k-1}}}{\partial u_{k-1}^{s_k}}Q(u)\Bigg|_{u=u_{(0)}}.$$

Theorem 2.2.1 *Let a vector function $g(\tau, z)$, $\tau \in R^{k-1}$, $z \in R^{k-1}$, possess the properties (i)–(iii). Then in a vicinity (say U) of the point $z_{(0)}$, there exists a vector function $\tilde{\tau} = \tilde{\tau}(z)$ such that the following hold:*

(I) $g(\tilde{\tau}(z), z) = 0$, $z \in U$.

(II) $\tilde{\tau}(z_{(0)}) = \tau_{(0)}$ and $\tilde{\tau}(z)$ is a real analytic vector function in U.

(III) The coefficients $\hat{\tau}_{(s)}$ of the expansion $\tilde{\tau}((z)$ into the Taylor series

$$\tilde{\tau}(z) = \sum_{s_1=0}^{\infty}\cdots\sum_{s_{k-1}=0}^{\infty}\tilde{\tau}_{(s)}(z_1 - z_{1(0)})^{s_1}\ldots(z_{k-1} - z_{k-1(0)})^{s_{k-1}}$$

can be calculated by recurrent formula that in the case $k = 2$ has the form

$$\tilde{\tau}_{(s+1)} = -J_{(0)}^{-1}g_{(s+1)}(\tilde{\tau}_{<s>}(z), z), s = 0, 1, \ldots,$$

where

$$\tilde{\tau}_{<s>}(z) = \tilde{\tau}_{(0)} + \sum_{j=1}^{s}\tilde{\tau}_{(j)}(z - z_{(0)})^j.$$

Note that assertions (I) and (II) are simply a reformulation of Theorem 1.8.1. Assertion (III) was established in Dette, Melas and Pepelyshev (2004b) and will be proved for the case of arbitrary k in Section 2.6.

Let us now apply this theorem to the function $g(\tau, z)$ given by relations (2.12). As is well known, the exponentials are real analytic at R^1 since

$$e^{-\lambda t} = 1 - \lambda t + \frac{(-\lambda t)^2}{2!} + \ldots + \frac{(-\lambda t)^n}{n!} + \ldots$$

and the series is convergent for any λ and t.

Additionally, multiplications and sums of real analytic functions are real analytic and, therefore,

$$\det(\psi_l(x_j))_{l,j=1}^m$$

is a real analytic function in $\Lambda = (\lambda_1, \ldots, \lambda_k)^T$ and $(x_1, \ldots, x_m)$ in R^{k+m}.

Note that the function $\varphi(\tau, \Lambda)$ and the vector function $g(\tau, z)$ are real analytic in a vicinity of the points $(\tau_{(0)}, \Lambda_{(0)})$, and $(\tau_{(0)}, z_{(0)})$, respectively, where $\Lambda_{(0)} = (1, \ldots, 1)$, $\tau_{(0)} = \tau^*(\Lambda_{(0)}) = (\gamma_1, \ldots, \gamma_{n-1})$, and $Z_{(0)} = (0, \ldots, 0)$,

In fact,

$$V(\xi_\tau, \Lambda) = \frac{\det M(\xi_\tau, \Lambda)}{\prod_{i<j}(\lambda_i - \lambda_j)^8}$$

and it can be verified [see Melas (1978)] that this function, codefined with preserving the continuity in points Λ such that $\lambda_i = \lambda_j$ for some $i \neq j$, is real analytic for arbitrary $\tau \in R^{m-1}$ and arbitrary $\Lambda \in R^k$.

Additionally, the function

$$\varphi(\tau, \lambda) = (V(\xi_\tau, \Lambda))^{1/m}$$

is real analytic as a rational degree of the real analytic function. It follows from here by the standard arguments that the function $\bar{\varphi}(\tau, z)$ and the vector function $g(\tau, z)$ are also real analytic for arbitrary $\tau \in R^{m-1}$ and $z \in R^{k-1}$.

Let us now calculate the matrix

$$\begin{aligned} J_{(0)} &= \left(\frac{\partial g_i(\tau, z)}{\partial \tau_j} \right)_{i,j=1}^{m-1} \Bigg|_{\tau=\tau_{(0)}, z=z_{(0)}} \\ &= \left(\frac{\partial^2}{\partial \tau_i \partial \tau_j} \bar{\varphi}(\tau, z) \right)_{i,j=1}^{m-1} \Bigg|_{\tau=\tau_{(0)}, z=z_{(0)}} . \end{aligned}$$

Due to (2.11) and the definition of $\bar{\varphi}(\tau, z)$ given in (2.12), we have

$$m(\varphi(\tau, z_{(0)})^m = e^{-2\sum_{i=1}^{m-1} \tau_i} \left(\prod_{i=1}^{m-1} \tau_i^2 \right) \prod_{i<j}^{m-1} (\tau_i - \tau_j)^2.$$

A direct calculation shows that

$$\frac{\partial \bar{\varphi}(\tau, Z_{(0)})}{\partial \tau_j} = \left[-1 + \frac{1}{\tau_j} + \sum_{s \neq j} \frac{1}{\tau_j - \tau_s} \right] \bar{\varphi}(\tau, z_{(0)}),\ i = 1, \ldots, m-1,$$

and the derivatives are equal to zero with $\tau = \tau_{(0)}$ by the definition of the point $\tau_{(0)}$.

Therefore,

$$(J_{(0)})_{ij} = \left. \frac{\partial^2 \bar{\varphi}(\tau, z_{(0)})}{\partial \tau_i \partial \tau_j} \right|_{\tau = \tau_{(0)}} = \frac{\bar{\varphi}(\tau_{(0)}, z_{(0)})}{(\gamma_j - \gamma_i)^2}\ (i \neq j),$$

$$(J_{(0)})_{ij} = \left. \frac{\partial^2 \bar{\varphi}(\tau, z_{(0)})}{\partial^2 \tau_j} \right|_{\tau = \tau_{(0)}} = - \left(\frac{1}{\gamma^2} + \sum_{s \neq j} \frac{1}{(\gamma_j - \gamma_s)^2} \right) \bar{\varphi}(\tau_{(0)}, z_{(0)}),$$

$i, j = 1, \ldots, m-1$.

Thus, for the matrix $J = J_{(0)}$, we have

$$(J)_{ij} > 0,\ i \neq j,\ J_{ij} < 0,\ i, j = 1, \ldots, m-1,$$

$$\sum_{j=1}^{m-1} (J)_{ij} = -\frac{\bar{\varphi}(\tau_{(0)}, z_{(0)})}{\gamma_j^2} < 0,\ i = 1, \ldots, m-1.$$

Due to the Hadamard criterion (see, e.g., Gantmacher (1998)), for an $(m-1) \times (m-1)$ matrix A

$$\det A \neq 0 \text{ if } (A)_{ii} > \sum_{i \neq j} |A_{ij}|,\ i = 1, 2, \ldots, m-1.$$

The matrix $(-J_{(0)})$ satisfies these conditions and, therefore, $\det J_{(0)} \neq 0$.

Thus, we proved that the function $g(\tau, z)$ determined by equalities (2.11) satisfies the conditions of Theorem 2.2.1 with $z_{(0)} = (0, \ldots, 0)$. $\tau_{(0)} = (\gamma_1 \ldots, \gamma_{m-1})$.

Consider now the case $k = 2$. In this case, the regression function is

$$\eta(x, \theta) = a_1 e^{-\lambda_i x} + a_2 e^{-\lambda_2 x},\ a_1, a_2 \neq 0, \lambda_1 \neq \lambda_2,$$

where $\lambda_1 > 0$ and $\lambda_2 > 0$. As will be shown in Chapter 6, LDMS designs are in this case LD among all (approximate) designs. Support points of these designs, as was already shown, do not depend on a_1 and a_2 and if λ_1 and λ_2 are multiplied by the same number $h > 0$, then the points should be divided by this number. Therefore, it will do to consider Λ such that $\lambda_1 + \lambda_2 = 2$ and to study the dependence of the support points of LDMS design on the parameter

$$z = z_1 = 1 - \lambda_1 = (\lambda_2 - \lambda_1)/2.$$

Let $z_{(0)} = 0$ and $\tau_{(0)} = (\gamma_1, \gamma_2, \gamma_3) = (0.467,\ 1.652,\ 3.879)$.

Note that the function $\varphi(\tau, z)$ is even,

$$\varphi(\tau, z) = \varphi(\tau, -z).$$

By this reason $\tau^*(z) = \tau^*(-z)$ and all odd coefficients $\tau^*_{(2j+1)}$, $j = 0, 1, \ldots$ are equal to zero. Therefore

$$\tau^*(z) = \tau_{(0)} + \sum_{t=1}^{\infty} \tau^*_{(2k)} z^{2t}. \tag{2.14}$$

The coefficients can be calculated by recurrent formulas of Theorem 2.2.1. These calculations can be easily realized with the help of the software package Maple. Some details of the implementation of the package are given in the Appendix of the present book.

First even coefficients calculated in this way are presented in Table 2.1.

Table 2.1: Coefficients $\tau_{<2t>}$, $t = 0, 1, \ldots, 6$

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

The method allows one to calculate as many coefficients as we like. Since the coefficients are already obtained, one can construct the corresponding designs simply by several first coefficients in the expansion (2.14).

However, we have a few problems here. The first problem concerns the radius of convergency of the series (2.14). Note that $0 \le |z| \le 1$ since

$$z = (\lambda_1 - \lambda_2)/2 \text{ and } (\lambda_1 + \lambda_2)/2 = 1.$$

Numerical studies show that the series are convergent for any $|z| < 1$. However, a strong theoretical proof of this fact is not obtained up to now.

The next problem consists of the determination of how many coefficients should be used in order to calculate support points of LDMS designs with an appropriate precision.

Denote $\tau(z, s) = \tau_{(0)} + \sum_{t=1}^{s} \tau_{2t} z^{2t}$ and

$$I_{(s)} = I_{(s)}(z) = \left(\frac{\det M(\xi_{\tau(z,s)}, z)}{\det M(\xi^*(z), z)} \right)^{1/m}, \ s = 0, 1, \ldots,$$

where $\xi^*(z) = \xi_{\tau^*(z)}$ is a LDMS design.

The value $I_{(s)}(z)$ is the efficiency of the design $\xi_{\tau(z,s)}$ constructed by s first even coefficients with respect to the LDMS for a given z. This value can be evaluated with the help of Kiefer's inequality (see Section 1.6)

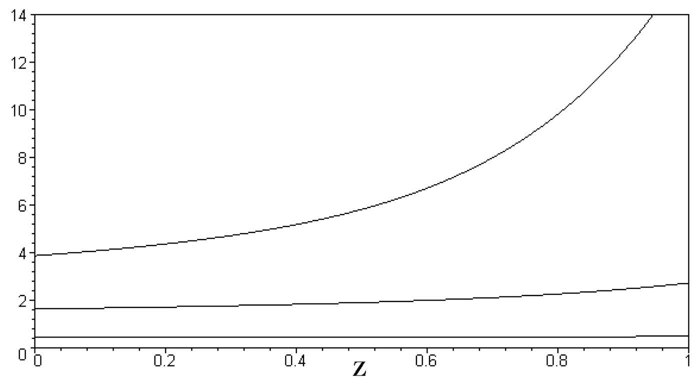

Figure 2.1: The dependence of support points of the LDMS designs on z for the exponential model with $k = 2$

without calculating the LDMS design. Some numerical results are given in Table 2.2. They represent an evaluation of $I_{(s)}(z)$, obtained with the help of Kiefer's inequality. Note that with $0 < z < 0.5$, $I_{(0)} = 1.00$ and there is no reason to calculate more coefficients.

Table 2.2: 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

From Table 2.2 we can conclude that with $|z| \leq 0.7$, it will do to use only one or two nonzero coefficients. However, for $z = 0.9$, we need 20 coefficients in order to obtain the efficiency greater than 0.995. Table 2.2 also shows that with $|z| \leq 0.9$, the expansions allow one to construct locally optimal designs with a very high precision. For $|z| > 0.9$, we can use a similar expansion with $z_{(0)} = 0,9$ as the initial point. The dependence of support points of the LD designs on z is presented in Figure 2.1. Note that we used 10 nonzero Taylor coefficients in order to construct this figure.

The next important question is: How efficient are LD designs with respect to equidistant designs usually implemented in practice?

Table 2.3: The efficiency of LD designs in respect to the best equidistant design

λ_1	1.1	1.3	1.5	1.7	1.9	1.95
λ_2	0.9	0.7	0.5	0.3	0.1	0.05
I	2.13	2.06	1.92	1.70	1.80	2.20

Denote by

$$\xi_{N,T} = \begin{pmatrix} 0 & T/(N-1) & \dots & T(N-2)/(N-1) & T \\ \frac{1}{N} & \frac{1}{N} & \dots & \frac{1}{N} & \frac{1}{N} \end{pmatrix}$$

the design located in N equidistant points at the interval $[0, T]$. For large N, the quality of this design is not very sensitive to the value of N, but it depends on T.

Consider the efficiency of LD designs constructed above with respect to equidistant design with an optimal choice of T; that is, we will take the value of T in such a way that the minimal efficiency of the equidistant designs for $z \in [0.1, 0.9]$ is the maximal one.

Our numerical results are given in Table 2.3. In this table, $T = 10$, $N = 20$;

$$I = \left(\frac{\det M(\xi_{\tau^*(\Lambda)}, \Lambda)}{\det M(\xi_{N,T}, \Lambda)} \right)^{1/m}, \; m = 4.$$

We see from Table 2.3 that in the most cases the efficiency of the LD design with respect to the equidistant design is more than 2 or close to 2. This means that the number of experiments in accordance with a LD design needed in order to achieve a given accuracy is approximately twice less than the same number for the best equidistant design if $\Lambda^{(0)} = \Lambda^*$. However, since Λ^* is unknown, these results describe the efficiency of LD designs only in an asymptotical sense. The influence of the choice of $\Lambda^{(0)}$ on the quality of LD designs can be studied numerically. However, in the following subsection we will show that the application of the functional approach can be used for such a study and allows one to compare LD designs with the maximin efficient ones.

2.2.3 Maximin efficient designs

Assume that it is known that $\Lambda^* \in \Omega$, where Ω is a given bounded and closed set in $R_+^k = \{\Lambda = (\lambda_1, \dots \lambda_k);\, \lambda_i > 0,\, i = 1, \dots, k\}$. Then a natural criterion of the efficiency of a given design is the value

$$\min_{\Lambda \in \Omega} \left(\frac{V(\xi, \Lambda)}{V(\xi(\Lambda), \Lambda)} \right)^{1/m}, \tag{2.15}$$

where $\xi(\Lambda)$ is a LD design, and for Λ such that $\lambda_i \neq \lambda_j$ $(i \neq j)$, the value $V(\xi,\Lambda)/V(\xi(\Lambda),\Lambda)$ is equal to $\det M(\xi,\Lambda)/\det M(\xi(\Lambda),\Lambda)$ [see the end of Section 2.2.1 for a discussion on this matter]. The value (2.15) will be called the minimal efficiency and the designs that maximize this value will be called maximin efficient D-optimal designs or, briefly, MME designs.

We will study the MME designs for the exponential model (2.1) and the set $\Omega = \Omega(\delta)$,

$$\Omega(\delta) = \Omega(\delta, c) = \{\Lambda = (\lambda_1, \ldots, \lambda_k) : (1-\delta)x_i \le \lambda_i \le (1+\delta)c_i,\ i = 1, \ldots, k\},$$

where $\delta \in (0,1)$, $c = (c_1, \ldots, c_k)$, $c_i > 0$, and $i = 1, \ldots, k$.

Let us restrict our attention to designs with the minimal support. In the following, it will be shown (see Theorems 2.2.2 and 2.2.3 and numerical results) that MME designs have the minimal support for sufficiently small δ and arbitrary c.

We have already proved that

$$\det M(\xi_\Delta, \Lambda) < \det M(\xi, \Lambda),$$

where

$$\xi_\Delta = \begin{pmatrix} \Delta & x_2 + \Delta & \ldots & x_m + \Delta \\ \frac{1}{m} & \frac{1}{m} & \ldots & \frac{1}{m} \end{pmatrix},\ \Delta > 0,$$
$$\xi = \begin{pmatrix} 0 & x_2 & \ldots & x_m \\ \frac{1}{m} & \frac{1}{m} & \ldots & \frac{1}{m} \end{pmatrix}.$$

Therefore, MME designs with a minimal support have the from

$$\xi_\tau = \begin{pmatrix} 0 & x_2 & \ldots & x_m \\ \frac{1}{m} & \frac{1}{m} & \ldots & \frac{1}{m} \end{pmatrix},\ \tau = (\tau_1, \ldots, \tau_{m-1}) = (x_2, \ldots, x_m). \quad (2.16)$$

Let us introduce the function

$$\hat{\varphi}(\tau, \Lambda) = \left(\frac{V(\xi_\tau, \Lambda)}{V(\xi_{\tau*(\Lambda)}, \Lambda)} \right)^{1/m},$$

where $\xi_{\tau^*(\Lambda)}$ is a LDMS design.

Theoretical studies (see Theorems 2.2.2 and 2.2.3) show that for sufficiently small $\delta > 0$,

$$\min_{\Lambda \in \Omega(\delta,c)} \hat{\varphi}(\tau, \Lambda) = \min\{\hat{\varphi}(\tau, (1-\delta)c), \hat{\varphi}(\tau, (1+\delta)c\}$$
$$= \min_{0 \le \alpha \le 1} \alpha \hat{\varphi}(\tau, (1-\delta)c) + (1-\alpha)\hat{\varphi}(\tau, (1+\delta)c).$$

Based on this, let us introduce the following class of designs. Let us say that a design is a maximin efficient design with a minimal structure or, briefly, MMEMS design, if this design is of the form (2.16), where $\tau = \hat{\tau}$ and $\hat{\tau}$ maximizes the value

$$\min_{0 \le \alpha \le 1} \alpha \hat{\varphi}(\tau, (1-\delta)c) + (1-\alpha)\hat{\varphi}(\tau, (1+\delta)c)$$

at the set of all vectors τ with positive coordinates.

In the case when intervals of possible values are the same for all parameters λ_i, $i = 1, \ldots, k$ (i.e., $c_1 = c_2 = \cdots = c_k$) the MMEMS designs can be found explicitly.

In order to describe these designs, let us denote

$$u = (\tau, \alpha) = (\tau_1, \ldots, \tau_{m-1}, \alpha),$$

$$\Phi(u, \delta) = \alpha \hat{\varphi}(\tau, (1-\delta)c) + (1-\alpha)\hat{\varphi}(\tau, (1+\delta)c),$$

$$\hat{\xi} = \xi_{\hat{\tau}} \text{ - MMEMS design.}$$

Let $\gamma_1, \ldots, \gamma_{m-1}$ be, as above, the roots of Laugerre's polynomial of degree $m-1$ with the associated parameter 1,

$$h = h(\delta) = 2\delta / \ln\left(\tfrac{1+\delta}{1-\delta}\right),$$

$$I(\delta) = [h(\delta) e^{(1-h(\delta))}]^{m(m-1)/2},$$

$$H = \left(\det M(\xi_{\tau^*(c)}, c)\right)^{1/m}.$$

Remember that $\tau^*(\gamma c) = \tau^*(c)/\gamma$ for any $\gamma > 0$. Also, it follows from here that

$$\hat{\varphi}(\tau, (1-\delta)c) = (\det M(\xi_\tau, (1-\delta)c))^{1/m} / (H(1-\delta)),$$

$$\hat{\varphi}(\tau, (1+\delta)c) = (\det M(\xi_\tau, (1+\delta)c))^{1/m} / (H(1+\delta)).$$

This simplifies theoretical and numerical studies of the MMEMS designs.

An explicit solution of the problem in the case $c_1 = c_2 = \cdots = c_k$ is given by the following theorem.

Theorem 2.2.2 *Consider model (2.1) and the set $\Omega = \Omega(\delta, c)$ of the form (2.4), where $c_1 = \cdots = c_k$. In this case the following hold:*

(I) There exists a unique MMEMS design for any fixed $c_1 > 0$ and $\delta < 1$. This design is

$$\hat{\xi} = \xi_{\hat{\tau}},\ \hat{\tau} = (\hat{\tau}_1, \ldots, \hat{\tau}_{m-1}),$$

$$\hat{\tau}_i = \gamma_i / (c_1 h(\delta)),\ i = 1, \ldots, m-1,$$

and

$$\Phi(\hat{u}, \delta) = I(\delta).$$

(II) This design is a locally D-optimal design for $\Lambda = c/h(\delta)$.

(III) For any sufficiently small positive δ, this design is MME among all (approximate) designs and its minimal efficiency is equal to $I(\delta)$.

Proof. Note that for $\Lambda = (c_1, \ldots, c_k)$, $c_1 = \cdots = c_k$ the value of $V(\xi, \Lambda)$ coincides with the value of determinant of the information matrix for the linear (by parameters) regression function

$$e^{-c_1 x}(\beta_1 + \beta_2 + \cdots + \beta_m x^{m-1}),$$

where $\beta = (\beta_1, \ldots, \beta_m)$ is the vector of estimating parameters and c_1 is a given number (as we already mentioned a detailed proof can be found in Melas (1978)). For this reason assertions (I) and (II) follows immediately from the results of Dette, Haines and Imhof (2003). Assertion (III) is a special case of Theorem 2.2.3(II). ∎

Note that the set of δ values for which assertion (III) holds can be found numerically. In particular, we found in such a way that assertion (III) is true for $k = 1$ with $\delta \leq 0.54$, for $k = 2$ it holds with $\delta \leq 0.22$, and for $k = 3$ it holds with $\delta \leq 0.18$. Thus, under realistic values of δ in the case $c_1 = \cdots = c_k$, MMEMS designs described in Theorem 2.2.2 are in fact MME designs among all (approximate) designs. It is also worth mentioning that in all of the cases, mentioned above, the minimal efficiency proves to be grater than 0.9, which can be easily checked by the explicit formula for $I(\delta)$.

In the case of arbitrary values $c_1, \ldots, c_k$, it seems does not possible to find MMEMS designs explicitly. However, the dependence of such designs on δ with a given c can be investigated with the help of constructing Taylor series in a way very similar to that was already applied to LDMS designs.

As is well known, the function of minimum is continuous. Also, we have already shown that the value $V(\xi_\tau, \Lambda)$ tends to zero with $\tau_{m-1} \to \infty$. Therefore, the function

$$\min_{0 \leq \alpha \leq 1} \Phi(u, \delta),\ u = (\tau, \alpha) \tag{2.17}$$

is bounded with $\tau \in R_+^{m-1}$ and there exists an MMEMS design (i.e., the design that maximizes (2.17) by $\tau \in R_+^{m-1}$).

Consider the equation system

$$\frac{\partial}{\partial u_i} \Phi(u, \delta) = 0,\ i = 1, \ldots, m. \tag{2.18}$$

Let $\hat{J}(\delta)$ be the Jacobi matrix of this system,

$$\hat{J}(\delta) = \left(\frac{\partial^2}{\partial u_i \partial u_j} \Phi(u, \delta) \right)_{i,j=1}^{m} \Bigg|_{u=u(\delta)},$$

where $u(\delta)$ is a solution of (2.18); the existence of this solution is provided by the following theorem.

Theorem 2.2.3 *Consider the regression model (2.1) for the set $\Omega = \Omega(c,\delta)$ defined in (2.4); the following assertions take place:*

(I) There exists a unique MMEMS design. Moreover, there exists a unique solution of the equation system (2.18), first $m-1$ components of this solution generate the vector $\hat{\tau}$, and the matrix $J(\delta)$ is invertible. The solution is a real analytic function of δ.

(II) If in a vicinity of $\Lambda = c$ the unique LDSM design is locally D-optimal among all (approximate) designs, then the MMEMS design is MME among all (approximate) designs for sufficiently small positive δ.

A proof of this theorem will be given in Section 2.6.3.

Note that as in the case of Theorem 2.2.2(III), the set of δ values for which assertion (II) is valid can be found numerically. For example, with $k = 2$ and $c = (1, 5)$, the MMEMS designs prove to be MME among all designs for all $\delta \leq 0.27$.

Theorem 2.2.3 justifies studying MMEMS and MME designs along the following steps:

1. Find numerically the MMEMS design for some value $\delta = \delta_0$ (in our calculations, we took $\delta_0 = 0.5$).

2. With the help of the recurrent formulas, construct the Taylor expansions for functions $\hat{\alpha}(\delta)$ and $\hat{\tau}_1(\delta), \ldots, \hat{\tau}_{m-1}(\delta)$.

3. Check whether the designs constructed are MME designs among all approximate designs for different values of δ by the equivalence theorem from Dette, Haines and Imhof (2003).

Let us illustrate the approach by examples.
With $k = 1$, the MMEMS designs are given by Theorem 2.2.2:

$$\hat{\xi} = \xi_{\hat{\tau}} = \begin{pmatrix} 0 & \hat{\tau}_1 \\ 1/2 & 1/2 \end{pmatrix},$$

where $\hat{\tau}_1 = 1/(c_1 h(\delta))$. A numerical calculation shows that this design is MME among all approximate designs if $\delta \leq 0.54$.

Let $k = 2$ and the set Ω be

$$\Omega = \Omega(z) = \{(\lambda_1, \lambda_2);\, c_i(1-\delta) \leq \lambda_i \leq c_i(1+\delta),\, i = 1, 2\}.$$

Without loss of generality, assume that $1 = c_1 \leq c_2$. Set $c_2 = 5$ (for other cases we obtain similar results).

Taylor coefficients for the functions $\hat{x}_i(\delta)$, $i = 2, 3, 4$ and $\hat{\alpha}(\delta)$ in a vicinity of $\delta = \delta_0 = 0.5$ are given in Table 2.4. Note that the series are convergent for $\delta \in [0, 1)$, and with $\delta < 0.8$, we need only three first coefficients to calculate MMEMS with a good precision. The values of the functions received by usage of the first 11 coefficients are depicted at Figure2.2. Note that the minimum efficiency is always achieved at the two points $(1-\delta)c$ and $(1+\delta)c$. The behavior of the function $\varphi(\hat{\tau}, \Lambda)$ with $\hat{\tau} = \hat{\tau}(0.5)$, $\Lambda \in \Omega(0.5)$ is shown in Figure 2.2. The dependence of the minimal efficiency on δ is presented in Figure 2.2.

Table 2.4: Coefficients in the Taylor expansion for the functions $\hat{x}_2(\delta)$, $\hat{x}_3(\delta)$, $\hat{x}_4(\delta)$, and $\alpha(\delta)$ by degrees of $(\delta - 0.5)$.

j	$\hat{x}_2$	$\hat{x}_3$	$\hat{x}_4$	α
0	0.17	0.69	2.06	0.44
1	0.05	0.34	1.16	−0.14
2	0.09	0.67	2.13	−0.09
3	0.11	0.99	2.87	−0.14
4	0.17	1.73	4.76	−0.19
5	0.28	2.99	7.85	−0.32

j	$\hat{x}_2$	$\hat{x}_3$	$\hat{x}_4$	α
6	0.49	5.33	13.56	−0.54
7	0.88	9.55	23.80	−0.94
8	1.60	17.26	42.59	−1.66
9	2.95	31.35	77.19	−2.97
10	5.50	57.22	141.40	−5.37
11	10.32	104.91	261.17	−9.74

The verification by the equivalence theorem mentioned above shows that the MMEMS designs are MME among all approximate designs with $\delta \leq 0.28$. Additionally, our calculations (not presented here) show that MMEMS designs have the minimal efficiency at 40–50% more than the best equidistant designs (such designs are often used in practice).

However, for $\delta > 0.28$, it is possible to construct even more efficient designs. For example, with $\delta = 0.5$ we constructed numerically a design that is MME among all approximate designs. This design has six support points and is approximately equal to

$$\begin{pmatrix} 0 & 0.140 & 0.440 & 1.048 & 1.75 & 3.25 \\ 0.24 & 0.18 & 0.19 & 0.16 & 0.13 & 0.10 \end{pmatrix}.$$

The minimal efficiency of this design is equal to 0.8431, whereas such efficiency for the MMEMS design is 0.7045. Note that for the LD design at the central point $\Lambda = (1, 5)$, this value is 0.6150, and for the best equidistant design, it is 0.5904.

For model (2.1) with three exponentials, we obtained similar results. However, the critical value of δ, for which the MMEMS designs remains MME among all designs, is smaller than that for the two exponential models.

2.3 Description of the Model

In this section we will introduce assumptions on the regression functions providing the application of the functional approach. The corresponding class of nonlinear regression models includes, in particular, the exponential models, considered in Section 2.2, as well as rational models and the three

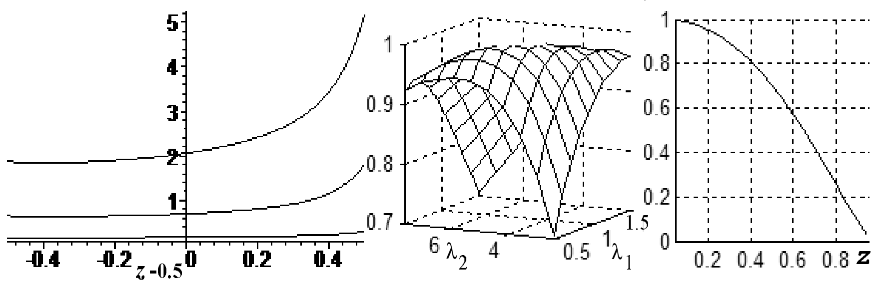

Figure 2.2: Functions $\hat{x}_2(\delta)$, $\hat{x}_3(\delta)$, and $\hat{x}_4(\delta)$; $z = \delta$(top) and the minimal of efficiency of MMEMS design (bottom right) with $\delta \in (0, 1)$. The efficiency of MMEMS design with $\delta = 0.5$ over $\Omega(0.5)$ (bottom left).

parameters logistic model. One more example is the Monod model to be studied in Chapter 8. For this class of models we introduce the basic equation determining the support points of locally D-optimal designs as implicit functions of values of the model parameters.

2.3.1 Assumptions and notation

Let us consider the general nonlinear regression model

$$y_j = \eta(x_j, \Theta) + \varepsilon_j,\ j = 1, \ldots, N, \tag{2.19}$$

where $y_1, \ldots, y_N \in R^1$ are experimental results, $\Theta = (\theta_1, \ldots, \theta_m)^T$ is the vector of unknown parameters, $\eta(x, \Theta)$ is a function of known form continuously differentiable along the parameters, $x_j \in \mathfrak{X}$, $\mathfrak{X}$ is a given set, and $\varepsilon_1, \ldots, \varepsilon_N$ are independent and identically distributed random values with zero expectation and a finite (unknown) variance $\sigma^2 > 0$.

Let us introduce the following notation (it was already given in Sections 1.7 and 2.2 but will be represented here for the sake of convenience of the reader):

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

$$f(x, \Theta) = (f_1(x, \Theta), \ldots, f_m(x, \Theta))^T,$$

$$M(\xi, \Theta) = \textstyle\int f(x, \Theta) f^T(x, \Theta) \xi(dx),$$

the information matrix,

$$\xi = \begin{pmatrix} x_1 & \ldots & x_n \\ \omega_1 & \ldots & \omega_n \end{pmatrix},\ x_i \neq x_j\ (i \neq j),\ x_i \in \mathfrak{X},\ \omega_i > 0,\ \sum \omega_i = 1,$$

approximate experimental design.

Denote by Θ^* the proper vector of the parameters. Designs maximizing the determinant of the information matrix for a fixed vector Θ will be called the LD designs. Usually, such designs depend only on a part of parameters (see Section 2.2). Without loss of generality, assume that these parameters are $\theta_{k+1}, \ldots, \theta_m$ and call them *nonlinear parameters*. Denote $\Theta_1 = (\theta_1, \ldots, \theta_k)^T$ and $\Theta_2 = (\theta_{k+1}, \ldots, \theta_m)^T$. Let us fix Θ_1 and consider the matrix $M(\xi, \Theta_2) = M(\xi, \Theta)$.

2.3.2 The basic equation

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

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

Let us consider designs LD among designs with the minimal support (i.e., with $n = m$). We call them LDMS designs. They often prove to be LD among all approximate designs. As was shown in Section 2.2, for such designs $\omega_1 = \cdots, = \omega_m = 1/m$.

Let $\Theta_2 \in \Omega$, where Ω is a given open set of possible values of Θ_2^*.

Assume that LDMS designs under $\Theta_2 \in \Omega$ have a fixed type (n_1, n_2, n_3), $n_1 + n_2 + n_3 = m$. Consider the case $n_1 = 1$ and $n_3 = 0$ (for all other cases, we can proceed in a very similar way). In this case, we will define the vector τ and the design ξ_τ as follows

$$\tau = (x_2, \ldots, x_m) = (\tau_1, \ldots, \tau_{m-1}),$$

$$\xi_\tau = \begin{pmatrix} x_1 & x_2 & \ldots & x_m \\ 1/m & 1/m & \ldots & 1/m \end{pmatrix}, \; x_1 = a.$$

Assume that the set Ω contains r linearly independent vectors and there are no $r + 1$ linearly independent vectors belonging to Ω. For example, for

$$\Omega = \{(\theta_{k+1}, \ldots, \theta_m)^T \, : \, \theta_i > 0, \; \sum_{i=k+1}^{m} \theta_i = m - k\}$$

$r = m - k - 1$.

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 inverse function $Q^{-1}(z)$ at the set $Z = Q(\Omega)$ is well defined. As an example, we can point out the vector function

$$z_i = 1 - \theta_{k+i}, \; i = 1, \ldots, r; \; r = m - k - 1, \tag{2.20}$$

introduced in Section 2.2.

Denote $\Theta^T(z) = \left(\Theta_1^T, (Q^{-1}(z))^T\right)$. Let $\mathcal{N}$ be the set of all vectors $z \in Z = Q(\Omega)$ such that

$$\det M(\xi_\tau, \Theta(r)) = 0$$

for any $\tau \in [a,b]^{m-1}$. For the case of exponential models described in Section 2.2, we have $\mathcal{N} = Q(\bar{\Omega})$, where $\bar{\Omega}$ is the set of all vectors $\Theta_2 \in \Omega$ such that two or more coordinates coincide with each other and Q is given by (2.20).

Let us introduce the following definition.

Definition 2.3.1 A vector function

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

where

$$V = \{\tau = (\tau_1, \ldots, \tau_{m-1}) \,:\, a < \tau_1 < \cdots < \tau_{m-1} < b\}$$

will be called the optimal design function if for any $z \in Z \setminus \mathcal{N}$, the design $\xi_{\tau*(z)}$ is a LDMS design for $\Theta^T = (\Theta_1^T, \Theta_2^T(z))$, $\Theta_2(z) = Q^{-1}(z)$ and for any sequence $z_{(1)}, z_{(2)}, \ldots$ such that $z_{(i)} \in Z \setminus \mathcal{N}$, $z_i \to \bar{z} \in \mathcal{N}$, $i \to \infty$,

$$\lim_{i\to\infty} \tau^*(z_i) = \tau^*(\hat{z}).$$

This definition is given for the case $n_1 = 1$ and $n_3 = 0$. The modification for other design types seems to be obvious.

Let us define the function

$$\varphi(\tau, z) = [\det M(\xi_\tau, \Theta(z))]^{1/m}; \tag{2.21}$$

the degree $1/m$ is introduced in order to secure a local convexity in a vicinity of the extreme points.

Due to the above assumption for any fixed $z \in Z \setminus \mathcal{N}$, the maximal value of the function $\varphi(\tau, z)$ by $\tau \in [a,b]^{m-1}$ is achieved in V. Therefore, a necessary condition for ξ_τ to be an LDMS design consists of vanishing of the derivatives

$$\frac{\partial}{\partial \tau_i}\varphi(\tau, z) = 0,\ i = 1, \ldots, m-1. \tag{2.22}$$

Set

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

The equation system (2.22) can be now written in the form

$$g(\tau, z) = 0. \tag{2.23}$$

This equation will be called *the basic equation of the functional approach.* It allows one to reduce the LDMS designs problem to the analysis of implicit functions. Such an analysis will be performed in Section 2.4. Now we will describe a class of regression functions for which this equation has a unique solution.

2.3.3 The uniqueness and the analytical properties

Let Z, N, and Q be as described above. Let us introduce the following assumptions:

A1. The functions

$$f_i(x, \Theta(z)),\ i = 1, \dots, m,$$

are real analytic by the variables $\{x_1, z_1, \dots, z_r\}$ at $(a, b) \times Z$.

A2. For $\Theta_2 \in \Omega$ all LDMS designs have the same type (n_1, n_2, n_3). For certainty, we will consider the case $n_1 = 1$ and $n_3 = 0$. Denote $H(\tau) = \prod_{1 \le i \le j \le m}(x_i - x_j)^2$, $\tau = (x_2, \dots, x_m)$, $x_1 = a$.

A3. There exists an algebraic polynomial $\Psi(z)$ such that

$$\inf_{z \in Z \setminus \mathcal{N}} \inf_{\tau \in V} \frac{\varphi^m(\tau, z)}{\Psi(z) H(\tau)} > 0,$$
$$\sup_{z \in Z \setminus \mathcal{N}} \sup_{\tau \in V} \frac{\varphi^m(\tau, z)}{\Psi(z) H(\tau)} < \infty.$$

Note that if the closure of Z does not intersect $\mathcal{N}$, we can take $\Psi(z) \equiv 1$. In this case, the assumption A3 means simply that the functions

$$f_1(x, \Theta), \dots, f_m(x, \Theta)$$

generate an extended Chebyshev system of order m on $[a, b]$ (see Section 1.9 for the definition) for all $\Theta_z = (\Theta_1, \Theta_2)$, $\Theta_z \in \Omega$.

Note also that the exponential regression functions introduced in Section 2.2 possess this property and all other assumptions were justified in that section.

Let us codefine the function

$$\bar{\varphi}(\tau, z) = \frac{\varphi(\tau, z)}{(\Psi(z))^{1/m}} = \left[\frac{\det M(\xi_{tau}, \Theta(z))}{\Psi(z)}\right]^{1/m}$$

by continuity with $z \in \mathcal{N}$. This is possible due to assumption A3.

A4. There exists a vector $z_{(0)} \in Z$ such that the equation system

$$\frac{\partial}{\partial \tau_i} \bar{\varphi}(\tau, z_{(0)}) = 0,\ i = 1, \dots, m-1,$$

has a unique solution with $\tau \in V$.

In Section 2.2, we have shown that this assumption holds for the exponential models with $z_{(0)} = (0, \dots, 0)$.

Now, the basic theorem of the functional approach can be formulated in the following way.

Theorem 2.3.1 *Let assumptions A1–A4 be fulfilled. Then the following hold:*

(I) There exists a unique optimal design function $\tau^*(z) : Z \to V$. It is a real analytic vector function in Z.

(II) Taylor coefficients of this vector function can be calculated by recurrent formulas given in Section 2.4.

A proof of this theorem will be given in Section 2.6.

2.4 The Study of the Basic Equation

In this section we will study (2.23) for a vector function $g(\tau, z)$ of a general form not necessarily connected with studying optimal experimental designs. We will obtain results stronger than that of Theorem 2.2.1(I, II), namely we will prove that under certain conditions, the function $\tau(z)$ determined implicitly by this equation is unique.

2.4.1 Properties of implicit functions

Assume that m and r are arbitrary natural numbers, and $m \geq r$ and $m \geq z$. Let

$$\hat{V} = \{\tau = (\tau_1, \dots, \tau_{m-1})^T : a \leq \tau_1 \leq \dots \leq \tau_{m-1} \leq b\},$$

$$V = \{\tau = (\tau_1, \dots, \tau_{m-1})^T : a < \tau_1 < \dots < \tau_{m-1} < b\},$$

and Z be an open one-connected set in R^r.

Let $\varphi(\tau, z)$, $\tau \in \hat{V}$, $z \in Z$, be a function of a general form real analytic in $V \times Z$, and $\varphi(\tau, z) \geq 0$.

Consider the case when $\varphi(\tau, z) = 0$ for some 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 $\Psi(z) = 0$ for $z \in \mathcal{N}$ and the function

$$\bar{\varphi}(\tau, z) = \varphi(\tau, z)/\Psi(z)$$

can be codefined in points $z \in \mathcal{N}$ by continuity.

Let $\bar{\varphi}(\tau, z)$ be the function codefined in the points $z \in \mathcal{N}$ in this way. Assume that $\bar{\varphi}(\tau, z) > 0$, $\tau \in V$, $z \in Z$, and

$$\inf_{\tau \in V} \frac{(\bar{\varphi}(\tau, z))^m}{H(\tau)} > 0,$$

$$\sup_{\tau \in V} \frac{(\bar{\varphi}(\tau, z))^m}{H(\tau)} < \infty,$$

for any $z \in Z$, where

$$H(\tau) = \prod_{i=1}^{m-1} (\tau_i - a)^2 \prod_{1 \leq i \leq j \leq m-1} (\tau_i - \tau_j)^2.$$

Let us denote

$$
\begin{aligned}
&g(\tau, z) = (g_1(\tau, z), \ldots, g_{m-1}(\tau, z)), \\
&g_i(\tau, z) = \frac{\partial}{\partial \tau_i}\varphi(\tau, z), \\
&G(\tau, z) = \left(\frac{\partial^2}{\partial \tau_i \partial \tau_j}\varphi(\tau, z)\right)_{i,j=1}^{m-1}, \\
&\bar{g}(\tau, z) = g(\tau, z)/\Psi(z), \\
&\bar{G}(\tau, z) = G(\tau, z)/\Psi(z) = \left(\frac{\partial^2}{\partial \tau_i \partial \tau_j}\bar{\varphi}(\tau, z)\right)_{i,j=1}^{m-1}.
\end{aligned}
$$

Consider the equations

$$
\begin{aligned}
g(\tau, z) &= 0, \\
\bar{g}(\tau, z) &= 0,
\end{aligned}
\tag{2.24}
$$

$z \in Z$, $\tau \in V$. For $z \in Z \setminus \mathcal{N}$, these equations are equivalent to each other.

Let us introduce the following assumptions:

(a) There exists a point $z_{(0)} \in Z$ such that (2.24) has a unique solution belonging to V.

(b) For any point z and any solution $z = \tau(z)$ of (2.24),

$$\det \bar{G}(\tau, z)\Big|_{\tau=\tau(z)} \neq 0.$$

Theorem 2.4.1 *Let the assumptions formulated above be satisfied. Then there exists a unique vector function $\tau^*(z) : Z \to V$ such that*

$$\bar{g}(\tau^*(z), z) = 0.$$

This vector function is real analytic for $z \in Z$ and satisfies the equation

$$G(\tau^*(z), z)\tau'_{z_i}(z) = (g(\tau, z))'_{z_i}\Big|_{\tau=\tau^*(z)}, \; i = 1, \ldots, m-1.$$

Proof. Due to assumptions (a) and (b) and the Implicit Function Theorem (Theorem 1.8.1), there exists a vicinity of the point $z_{(0)}$ such that there exists a unique vector function, say $\bar{\tau}(z)$, satisfying (2.24). This vector function is real analytic. Let U be a union of all such vicinities. Then $\bar{\tau}(z)$ can be extended to U in a unique way and this extended function is real analytic in U.

Suppose that $U \neq Z$. Denote the closure of U by $\bar{U}$. Since $U \neq Z$, there exists a point $\bar{z} \in \bar{U} \setminus U$, $\bar{z} \in Z$. Then there exists a sequence $z_{(1)}, z_{(2)}, \ldots,$ such that $z_{(i)} \in U$ and $\lim_{i\to\infty} z_{(i)} = \bar{z}$. Denote by $\bar{\tau}$ the limit

$$\lim_{i\to\infty} \bar{\tau}(z_{(i)}).$$

Then we have

$$\bar{g}(\bar{\tau}, \bar{z}) = 0.$$

Suppose that $\bar{\tau} \in V$. Then, due to assumption (b),

$$\det \bar{G}(\bar{\tau}, \bar{z}) \neq 0$$

and there exists a vicinity of point $\bar{z}$, say W, and vector function $\tau_{(1)}(z)$ such that $\tau_{(1)}(\bar{z}) = \bar{\tau}$ and this vector function is real analytic in this vicinity. Moreover, for sufficiently large i, $z_{(i)}$ belongs to this vicinity. It follows from here that $\tau_{(1)}(z)$ and $\bar{\tau}(z)$ coincide in $W \cap Z \neq \emptyset$. Therefore, $\bar{\tau}_{(1)}(z)$ is a real analytical extension of $\bar{\tau}(z)$ to W and $W \bar{\subset} U$. This is impossible by our supposition and we obtained a contradiction.

Now, let $\bar{\tau} \in \hat{V} \setminus V$. Denote $\tau_{(i)} = \tau(z_{(i)})$, $i = 1, 2, \ldots$. Then

$$\begin{aligned} &\lim_{j\to\infty} \frac{\partial}{\partial \tau_i} \left[\frac{(\bar{\varphi}(\tau, z_j))^m}{Q(\tau)} \right] \Bigg|_{\tau=\tau_{(j)}} \\ &\quad = \lim_{j\to\infty} \left\{ \left\{ \frac{\partial}{\partial \tau_i} \left(\bar{\varphi}(\tau, z_{(1)})\right)^m \right\} \Bigg|_{tau=\tau_{(j)}} \Big/ Q(\tau_{(j)}) \right. \\ &\quad \left. - \left(\bar{\varphi}(\tau_{(j)}, z_{(i)})\right)^m \frac{\partial Q(\tau)/\partial \tau_i}{Q^2(\tau)} \Bigg|_{\tau=\tau_{(i)}} \right\} = \infty. \end{aligned}$$

However, due to our assumption, the function

$$\frac{\bar{\varphi}(\tau, z)}{Q(\tau)}$$

is real analytic in $V \times Z$, and the limit should be finite. The obtained contradiction shows that $U = Z$. In a similar way, it can be proved that for any $z \in Z$, (2.24) has a unique solution. ∎

In order to apply Theorem 2.4.1 to the function $\varphi(\tau, z)$ defined in Section 2.3, we need only to verify property (b). To this end we will introduce a representation for the Jacobi matrix of (2.24).

2.4.2 Jacobian of the basic equation

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

Let m, r, and t be arbitrary natural numbers, $T \subset \mathbb{R}^{m-1}, Z \subset \mathbb{R}^r$, and $\mathfrak{A} \subset \mathbb{R}^t$ be arbitrary open sets, and $\mathfrak{A}$ be 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

$$\varphi(\tau, z) = \min_{a \in \mathfrak{A}} q(\tau, a, z), \tag{2.25}$$

where the minimum is attained for any $\tau \in T$ and $z \in Z$. Since the function $q(\tau, a, z)$ is strictly 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.

For any fixed z, let there exist a point $\tilde{\tau} = \tilde{\tau}(z)$ satisfying the equation $\frac{\partial}{\partial \tau} \varphi(\tau, z) = 0$.

Consider the following matrices:

$$\begin{aligned} E &= \left(\tfrac{\partial^2}{\partial \tau_j \partial \tau_i} q(\tau, a, z) \right)_{i,j=1}^{m-1}, \\ B &= \left(\tfrac{\partial^2}{\partial \tau_j \partial a_i} q(\tau, a, z) \right)_{i,j=1}^{t,m-1}, \\ D &= \left(\tfrac{\partial^2}{\partial a_j \partial a_i} q(\tau, a, z) \right)_{i,j=1}^{t} \end{aligned} \tag{2.26}$$

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

Theorem 2.4.2 *Under the above conditions, the following formula is valid:*

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

Let us apply this theorem to the function $\varphi(\tau, z)$, defined by (2.21).

Denote the set of all positive definite $m \times m$ matrices $A = (a_{ij})$, such that $a_{mm} = 1$ by $\mathcal{A}$. Assign a number $\nu = \nu(i, j)$ in alphabetical order to each pair of indices (i, j), $i \le j$, $i, j = 1, \ldots, m$, where $(i, j) \ne (m, m)$. For any vector $a \in \mathbb{R}^t$, $t = m(m+1)/2 - 1$, define a matrix $A(a)$ that satisfies the following relations:

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

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

$$q(\tau, a, z) = (\det A(a))^{-1/m} \operatorname{tr} \left(A(a) M(\xi, z) \right) / m. \tag{2.27}$$

Consider the function $\varphi(\tau, z) = (\det M(\xi, z))^{1/m}$. It is known (Karlin and Studden, 1966, Chap. 10.2) that (2.25) is valid for this function. It can also be checked that the function (2.27) possesses the required properties. Therefore, by Theorem 2.4.2,

$$J(\tilde{\tau}(z), z) = E - B^T D^{-1} B, \tag{2.28}$$

where $\tilde{\tau}(z) = \tau^*(z)$. 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:

$$\begin{aligned}
&E = \operatorname{diag}\{E_{11}, \ldots, E_{m-1m-1}\},\\
&E_{ii} = \delta(a^*) \frac{\partial^2}{\partial x^2} (f^T(x) A(a^*) f(x))\Big|_{x=x^*_{i+1}}, \; i = 1, \ldots, m-1,\\
&A(a^*) = \text{const} \left(M(\xi_{\tau^*(z)}, z) \right)^{-1},\\
&B = (b_{\nu k})_{\nu,k=1}^{t,m-1},\\
&b_{\nu k} = 2\delta(a^*) \frac{\partial}{\partial x} (f_i(x) f_j(x))\Big|_{x=x^*_k}, \nu = \nu(i,j).
\end{aligned} \tag{2.29}$$

Remark 2.4.1 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 $J = E - B^T D^{-1} B$, J is negative definite if either of conditions (1) and (2) is valid.

This remark will be applied in Section 2.6.2 in order to prove that the matrix J is invertible under assumptions A1–A4.

2.4.3 On the representation of implicit functions

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 subsection 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 and Pepelyshev (2004b).

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

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

where $z_{(0)}$ is a given point.

Introduce also the notation

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

$t = 0, 1, \ldots$, and

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

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, and $\bar{\psi}(z)$ 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

$$I_t = U_{j=0}^t S_j,$$

$$\tau_{<I_t>}(z) = \textstyle\sum_{s \in I_t} \tau_{(s)} (z - z_{(0)})^s, \quad \tau_{(s)} = (\tau(z))_{(s)},$$

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

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 2.4.3 *Under conditions (a) and (b) for the function $\tau(z)$, defined in Theorem 2.4.1, the following formulas hold:*

$$(\tau(z))_{(s)} = -J_{(l)}^{-1} g(\tau_{<I>}(z), z)_{(s+l)}, \tag{2.30}$$

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

If 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 (2.30). 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

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

Let

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

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

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

Theorem 2.4.4 *With $p > 0$, Theorem 2.4.3 remains true with (2.30) replaced by*

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

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

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

2.4.4 The monotony property

Let us obtain another representation for the Jacobi matrix. It is based on the known formula for the derivative of the matrix determinant. This representation is to help us to derive the monotony of coordinates of the optimal design function for some forms of regression.

At first, let $x \in \mathfrak{X}$ and $z \in Z$, where Z is some bounded set in $\mathbb{R}^r$ and $f_i(x, z), i = 1, \ldots, m$, are arbitrary twice differentiable with respect to x functions.

Let the function $\varphi(\tau, z)$ be defined by (2.21). Let assumptions A1–A4 be satisfied. We will use the formula of differentiating the matrix determinant (see, e.g., Fedorov (1972)).

$$\frac{\partial}{\partial \alpha} \det M(\alpha) = \det M(\alpha) \left(\operatorname{tr} M^{-1}(\alpha) \frac{\partial}{\partial \alpha} M(\alpha) \right),$$

as well as the formula

$$\frac{\partial}{\partial \alpha} M^{-1}(\alpha) = -M^{-1}(\alpha) \frac{\partial M(\alpha)}{\partial \alpha} M^{-1}(\alpha)$$

and the explicit form of matrix $M(\xi, z)$:

$$M(\xi, z) = \sum_{i=1}^{m} f(x_i) f^T(x_i)/m,$$

where $f(x) = f(x, z)$ and

$$\xi = \begin{pmatrix} x_1 & \ldots & x_{m-1} & x_m \\ \frac{1}{m} & \ldots & \frac{1}{m} & \frac{1}{m} \end{pmatrix}.$$

Let us calculate the derivatives of the function

$$\varphi(\tau, z) = \left(\det M(\xi_\tau, z)\right)^{1/m},$$

$$\xi_\tau = \begin{pmatrix} \tau_1 & \dots & \tau_{m-1} & b \\ \frac{1}{m} & \dots & \frac{1}{m} & \frac{1}{m} \end{pmatrix}.$$

We obtain

$$\begin{aligned}\frac{\partial \varphi(\tau,z)}{\partial \tau_i} &= \frac{1}{m^2}\varphi(\tau,z)\,\mathrm{tr}\,M^{-1}(\xi,z)(f(\tau_i)f^T(\tau_i))' \\ &= \frac{2}{m^2}\varphi(\tau,z)f^T(\tau_i)M^{-1}(\xi,z)f'(\tau_i),\end{aligned}$$

$i = 1, \dots, m-1$. Let z be fixed and τ be such that

$$g(\tau,z) = \left(\frac{\partial}{\partial \tau_i}\varphi(\tau,z)\right)_{i=1}^{m-1} = 0. \tag{2.31}$$

Moreover, let τ be a local maximum of the function $\varphi(\tau,z)$. Set $F = (f_j(x_i))_{i,j=1}^m$. Then the relation $M = FF^T/m$ is valid,

$$\begin{aligned} f^T(\tau_i)M^{-1}(\xi,z)f(\tau_j) &= mf^T(\tau_i)(F^{-1})^TF^{-1}f(\tau_j) \\ &= m\, e_{i+1}^T e_{j+1}^T = \begin{cases} 0 & i \neq j \\ m & i = j \end{cases}. \end{aligned}$$

Let us consider the matrix

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

where $\tau(z)$ is the unique solution of (2.31).

Using these relations and the formula of the inverse matrix differentiation, derive

$$\begin{aligned}(G)_{ij} &= \frac{\partial^2}{\partial \tau_i \partial \tau_j}\varphi(\tau,z) \\ &= -\frac{4}{m^3}\varphi(\tau,z)\left(f^T(\tau_i)M^{-1}(\xi,z)f'(\tau_j)\right)\left(f^T(\tau_j)M^{-1}(\xi,z)f'(\tau_i)\right) \\ &= -\frac{4}{m}\varphi(\tau,z)\left(e_i^TF^{-1}f'(\tau_j)\right)\left(e_j^TF^{-1}f'(\tau_i)\right)\end{aligned}$$

for $i \neq j$, $i,j = 1,\dots,m-1$. For calculating the diagonal elements of the matrix, let us also use the following relation:

$$f^T(\tau_i)M^{-1}(\xi,z)f'(\tau_i) = \frac{m^2}{2\varphi(\tau,z)}\frac{\partial\varphi(\tau,z)}{\partial\tau_i} = 0,\ i = 1,\dots,m-1.$$

The direct differentiation gives the following result

$$\begin{aligned}(G)_{ii} &= \frac{\partial^2}{\partial \tau_i \partial \tau_i}\varphi(\tau,z) \\ &= \frac{2}{m^2}\varphi(\tau,z)f^T(\tau_i)M^{-1}(\xi,z)f''(\tau_i) \\ &= \frac{2}{m}\varphi(\tau,z)e_i^TF^{-1}f''(\tau_i), i = 1,\dots,m-1.\end{aligned}$$

Now, assume the that functions $f_i(x) = f_i(x,z), i = 1,\ldots,m$, form an ET-system (see Section 1.9 for the definition) of the first order under any fixed $z \in Z$.

Since the matrix F^{-1} is formed by the cofactors of the elements of matrix F, divided by its determinant, then, for $j > i$, $i > 2$, we have

$$e_i^T F^{-1} f'(\tau_j)$$

$$= \frac{\det\left(f(x_1) \vdots f(x_2) \vdots \ldots \vdots f(x_{i-1}) \vdots f'(x_j) \vdots f(x_{i+1}) \vdots \ldots \vdots f(x_m)\right)}{\det F}$$

(with the evident changes for $i \leq 2$). Inserting a column $f'(x_j)$ between a line $f(x_j)$ and the following one, derive

$$e_i^T F^{-1} f'(\tau_j) = (-1)^{j-i} \det \tilde{F} / \det F,$$

$$\tilde{F} = \left(f(x_1) \vdots f(x_2) \vdots \ldots \vdots f(x_{i-1}) \vdots f(x_{i+1}) \vdots \ldots \vdots f(x_j) \vdots f'(x_j) \vdots \ldots \vdots f(x_m)\right).$$

By definition of the ET-system of the first order, $\det \tilde{F} > 0$. Thus,

$$\operatorname{sign}\left[e_i^T F^{-1} f'(\tau_j)\right] = (-1)^{j-i}.$$

Similarly, for $i < j$, we have

$$\operatorname{sign}\left[e_{j+1}^T F^{-1} f'(\tau_i)\right] = (-1)^{i-j}.$$

Therefore, for $i \neq j$

$$\operatorname{sign}(G)_{ij} = (-1)(-1)^{j-i-1}(-1)^{i-j} = 1.$$

It will be proved in Section 2.6.2 that the matrix G is negative definite. Let us use the following statement (see Szegö, 1959): If matrix A is positive definite and each of its off-diagonal elements is negative, then all of the elements of the matrix A^{-1} are positive. Since the matrix G is negative definite and its off-diagonal elements are positive, then the matrix $A = -G$ possesses the required properties. Applying the above statement, we have that all of the elements of the matrix G^{-1} are negative.

Thus, we have derived the following result.

Lemma 2.4.1 *If functions $f_i(x,z), i = 1,\ldots,m, x \in \mathfrak{X}, z \in Z$, are twice continuously differentiable on $\mathfrak{X}$ and form an ET-system of the first order for any fixed $z \in Z$, the matrix G is invertible and all of the elements of matrix G^{-1} are negative.*

Let the conditions of Lemma 2.4.1 be satisfied. By Theorem 2.3.1, the optimal design function $\tau(z) : Z \to V$ is uniquely determined. Let L_j stand for the vector

$$\frac{\partial}{\partial z_j} g(\tau, z) = \left(\frac{\partial^2}{\partial \tau_i \partial z_j} \varphi(\tau, z)\right)_{i=1}^{s-u}, \; j = 1, 2, \ldots, r.$$

By the Implicit Function Theorem, we have

$$\tau_{z_j}^{'} = -G^{-1}L_j. \tag{2.32}$$

Thus, if all of the elements of vector L_j are positive, then

$$(\tau_i(z))_{z_j}^{'} < 0,\ i = 1, \dots, m-1;$$

that is, all of the coordinates of function $\tau(z)$ monotonously decrease with respect to z_j.

Let us introduce a class of regression functions for which all the unfixed points of a locally D-optimal design monotonously depend on each parameter. We will show further that this class contains the exponential models considered in Section 2.2, as well as some rational models.

Consider a real function $K(x, y)$, defined for $(x, y) \in \mathfrak{X} \times \mathfrak{X}_1$, where $\mathfrak{X}$ and $\mathfrak{X}_1$ are intervals. Let function $K(x, y)$ be an extended strictly positive kernel of the m-th order (ESP(m)) along both variables. The corresponding definition can be found in Karlin and Studden (1966, Chap. I).

Consider the regression function

$$\eta(x, \Theta) = \sum_{i=1}^{k} \theta_i K(x, \theta_{i+k}),\ \theta_i \neq 0, i = 1, \dots, k,\ m = 2k.$$

Let the functions $f_i(x, \Theta) = \frac{\partial}{\partial \theta_i} \eta(x, \Theta)$, $i = 1, \dots, m$, for $\Theta_2 = (\theta_{k+1}, \dots, \theta_m)^T = (z_1, \dots, z_k)^T \in Z \subset \mathfrak{X}_1^k$ be real analytic.

By the definition of ESP(m) functions $f_i(x, z), i = 1, \dots, m$ (for fixed Θ_1) at any fixed $z \in Z$ form an ET-system. Therefore, the optimal design function $\tau(z)$ is uniquely determined at $z \in Z$.

Assume that for some point $z_{(0)}$ for $z = z_{(0)}, \tau = \tau(z_{(0)})$ the following inequality is valid:

$$\frac{\partial^2 \varphi(\tau, z)}{\partial \tau_i \partial z_j} > 0,\ i = 1, \dots, s-u,\ j = 1, \dots, k. \tag{2.33}$$

Theorem 2.4.5 *Under the above conditions, all of the components of the vector function $\tau(z)$ decrease with respect to each of $z_1, \dots, z_k$ in a strictly monotonous way.*

Proof. By Lemma 2.4.1 and formula (2.32), it is sufficient to prove that

$$\frac{\partial^2 \varphi(\tau, z)}{\partial \tau_i \partial z_j} > 0,\ i = 1, \dots, m-1,\ j = 1, \dots, k.$$

for any $z \in Z$.

Let $\overline{z}_1 < \overline{z}_2 < \cdots < \overline{z}_k$. Set $v = z_j$ and consider the function $\partial\varphi(\tau, z)/\partial\tau_\nu$ $(\nu = 1, \dots, m-1)$ as a function of v under fixed z_i, $i = 1, \dots, k$,

$i \neq j$, $x_1, \ldots, x_{m-1}$. Denote this function by $h(v)$. Note that the function $h(v)$ has second-order zeros at points z_i, $i \neq j, i = 1, \ldots, k$, since the determinants corresponding to $h(v)$ and $h'(v)$ have common lines. Moreover, $h(\overline{z}_j) = 0$. Since $\varphi^m(\tau, z)$ equals

$$\begin{aligned} &\text{const}\det\left(K(x_j, z_1), K'_z(x_j, z_1), \ldots,\right. \\ &\left.K(x_j, z_k), K'_z(x_j, z_k)\right)_{j=1}^{m}, \end{aligned} \tag{2.34}$$

then the function $h(v)$ has no more than $2k - 1$ zeros (counting with their multiplicities). Therefore, $h'(\overline{z}_j) \neq 0$. The case that some $\overline{z}_i$ coincide with one another can be processed in a similar way (here determinant (2.34) is to be modified as is stated in the definition of the ESP kernel). Thus, the functions

$$\frac{\partial^2}{\partial \tau_i \partial z_j} \varphi(\tau, z)$$

do not vanish. Since, by assumption, condition (2.33) is valid for $z = z_{(0)}$, it is valid for any $z \in Z$. ∎

Consider two examples.

Example 2.4.1 Algebraic sum of simplest fractions.

Let

$$K(x, y) = \frac{1}{x + y},\ \mathfrak{X} \subset [0, \infty),\ \mathfrak{X}_1 \subset [0, \infty).$$

It is known that such a function is an ESP kernel of any order (Karlin and Studden 1966, Chap. I). The corresponding regression function takes the form

$$\eta(x, \Theta) = \sum_{i=1}^{k} \frac{\theta_i}{x + \theta_{i+k}},\ x \in [0, \infty),$$

$\theta_{i+k} > 0$, $\theta_i \neq 0$, $i = 1, \ldots, k$. For corresponding basis functions $f_i(x, \Theta), i = 1, \ldots, k$, condition (2.33) can be verified directly. It can be demonstrated also that condition A4 is satisfied. These models will be thoroughly investigated in Chapter 5.

Example 2.4.2 Algebraic sum of exponential functions.

Let

$$K(x, y) = e^{xy},\ \mathfrak{X} \subset (-\infty, \infty),\ \mathfrak{X}_1 \subset (-\infty, \infty).$$

This function is an ESP kernel of any order [Karlin and Studden (1966, Chap. I)]. The corresponding regression function takes the form

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

$\theta_i \neq 0, i = 1, \ldots, k.$,k

2.5 Three-Parameter Logistic Distribution

Consider the function

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

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

$$\eta(x, \Theta) = \frac{\theta_1 x^{\theta_2}}{\theta_3 + x^{\theta_2}}, \tag{2.35}$$

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

We will construct locally D-optimal designs for model (2.35) 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 obtain

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

where

$$\xi = \begin{pmatrix} x_1 & x_2 & x_3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix},\ \zeta_\xi = \begin{pmatrix} t_1 & t_2 & t_3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix},$$

$$t_i = x_i^{\theta_2},\ i = 1, 2, 3,$$

$$\bar{M}(\zeta_\xi, \theta_3) = \sum_{i=1}^{3} f(t_i, \theta_3) f^T(t_i, \theta_3)/3,$$

$$f(t, \theta) = \left(\frac{t}{\theta + t}, \frac{t}{(\theta + t)^2}, \frac{t \ln t}{(\theta + t)^2} \right)^T.$$

Set

$$z = 1/\theta_3,\ r = 1,\ \Omega = [0, \infty),\ \psi(z) = z^6,\ \mathcal{N} = \{0\}. \tag{2.36}$$

Assumption A1 follows here from the properties of elementary functions, A2 and A3 follows from the results of Dette, Melas, and Wong (2004b). It was also proved there that a locally D-optimal design has the type $(0, 2, 1)$ and is unique. It can be also proved that A4 holds for the considered model.

Thus, due to Theorem 2.3.1, 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 $0 \leq a < b, \theta_2$ optimal designs can be calculated by a scale transformation. With $\theta_3 \to \infty$ and $z = \frac{1}{\theta_3} \to 0$, we obtain

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

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 $x_1^*(0) = 0.15370$ and $x_2^*(0) = 0.61680$.

By the recurrent formulas (2.30) given in Section 2.4, we calculated the Taylor coefficients with $z_{(0)} = 0$. The first coefficients are represented in Table 2.5.

Table 2.5: 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

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

$$\begin{aligned} &\max_{x\in[0,1]} |d(x, \xi_{<n>}(z)) - 3| \le 10^{-5}, \\ &d(x, \xi) = f^T(x) M^{-1}(\xi, z) f(x), \end{aligned} \tag{2.37}$$

where

$$f(x) = \frac{\partial \eta(x, \Theta)}{\partial \theta_i}, \; M(\xi, z) := M(\xi, \Theta(z)), \; \Theta(z) = (1, 1, 1/z)^T.$$

Note that due to the Kiefer–Wolfowitz equivalence theorem (see Section 1.5), a design satisfying condition (2.37) will be very close to a locally D-optimal design. Numerical calculations show that $\bar{z}_{10} \approx 0.705$ and $\bar{z}_{20} \approx 0.865$.

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 2.6 and 2.7, respectively. It proves that for the first expansion with 20 coefficients, the inequality (2.37) holds with $0 < z \le 2.7$. For the second expansion with the same number of the coefficients, it holds for $0.6 \le z \le 13.8$.

The behavior of the design points for $0 \le z \le 10$ is presented in Figure 2.3. We used the first expansion for $z \le 1$ and the second for $1 \le z \le 10$ to construct Figure 2.3.

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

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

proves to be very high with $z \le 1$ ($\theta_3 \ge 1$). This efficiency is presented in Table 2.8.

Table 2.6: 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 2.7: 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

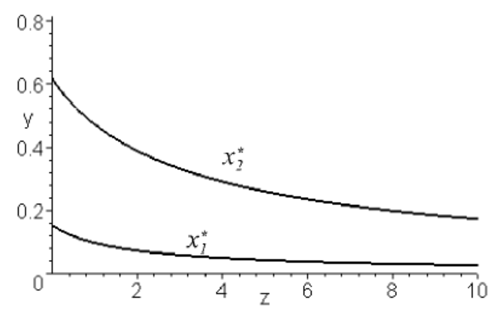

Figure 2.3: The dependence of the support points x_1 and x_2 on z

Table 2.8: Efficiency of designs $\xi(0)$ and $\xi(1)$ 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

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 2.8. Moreover, numerical calculations show that the design $\xi(z^*) = \xi_{\tau^*(z^*)}$ with $z^* = 0.5$ has a 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 is the design maximizing the minimum by $z \in [0.1, 1]$ of the efficiency among all (approximate) designs, was constructed numerically in Dette, Melas, and Wong (2004b). 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 and Wong (2004b) has four support points with unequal weights and its minimal efficiency equals 0.885. However, for example, design $\xi(1)$, the locally optimal design for z=1, has the minimal efficiency 0.5430 on $[1, 10]$. This design is rather bad! It requires almost twice as many observations as $\xi(4)$ to achieve the same accuracy of the estimates of the parameters if the true value of z equals 10.

Thus, we see that the approach allows very efficient calculation of locally D-optimal designs and gives an opportunity to study their efficiency.

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

2.6 Appendix: Proofs

We begin with the proofs for the theorems of Section 2.4.

2.6.1 Proof of Theorems 2.4.2, 2.4.3, and 2.4.4

Proof of Theorem 2.4.2. 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 $\tau = \tilde{\tau} = \tilde{\tau}(z)$ and $a = \tilde{a} = \tilde{a}(z, \tilde{\tau}(z))$.

Consider this vector equality at fixed z and arbitrary a and τ as an equation system that 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

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

An immediate calculation now gives

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

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 obtain

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

■

Proof of Theorem 2.4.3. 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 2.6.1 *Under condition (b) and with $p = 0$ and $l = 0$ the following equalities are valid:*

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

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

Proof of Lemma 2.6.1. 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 obtain for $t = 1, \ldots, K-1$:

$$\begin{aligned}
&\frac{\partial^t}{\partial z^t} g(\tau(z), z)|_{z=z_{(0)}} \\
&\quad = t! J_{(0)} \tau_{(t)} + \frac{\partial^t}{\partial z^t} g(\tau_{(0)}, z_{(0)}) + \cdots \\
&\quad + \sum_{i_1, \ldots, i_t = 1}^{m} \frac{\partial^t}{\partial \tau_{i_1} \ldots \partial \tau_{i_t}} g(\tau_{(0)}, z_{(0)}) \tau_{i_1(1)} \cdots \tau_{i_t(1)} i_1! \cdots i_t!,
\end{aligned} \tag{2.38}$$

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 the case $k > 1$, the proof is similar. ■

Return to the proof of Theorem 2.4.3. Let $k = 1$ and $l = 0$. Note that on the right-hand side of (2.38), 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 $z_{(0)}$,

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

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

Proof of Theorem 2.4.4. At first, consider $l = 0$. Note that

$$(g(\tau_{<I>}(z), z))_{(s+u)} = \sum_{w+v=s+u} a_{(w)} \tilde{g}(\tau_{<I>}(z), z)_{(v)} \tag{2.39}$$

for any collection of indexes I,

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

For $w = u$, vector s is the only vector v such that $w + v = s + u$. Let $s \in \hat{S}_n$, $I = \hat{I}_n$. Note that for $w \neq u$, any vector v such that $w + v = s + u$ belongs to set $\hat{S}_t$, $t \leq n-1$, from which it follows that the right-hand side of (2.39) has the form

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

It can be verified by direct calculation that $J_{(u)} = a_{(u)} \tilde{J}_{(0)}$. Therefore, Theorem 2.4.4 is valid at $l = 0$. For arbitrary l, its validity can be verified by direct calculation. ∎

2.6.2 Proof of Theorem 2.3.1

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 $\Theta^{0^T} = \left(\Theta_1^{0^T}, (q^{-1}(z))^T\right)$. This function should satisfy equation (2.10) and due to the Implicit Function Theorem (Gunning and Rossi, 1965) we need only to prove that the Jacobi matrix, J, is invertible. For this, it will do to prove that matrix B is of full rank. 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

$$\sum_{s=2}^{m} \left[f_i(x_s^*) f_j^{'}(x_s^*) + f_i^{'}(x_s^*) f_j(x_s^*)\right] d_s = 0, \tag{2.40}$$

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

Note that (2.40) holds also for $(i, j) = (m, m)$. In fact, since

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

is a saturated locally D-optimal design, we have

$$\frac{\partial}{\partial x_s} \det M(\xi_{\bar{\tau}}, z) = \sum_{i,j=1}^{m} (f_i(x_s^*) f_j(x_s^*))' d_{ij} = 0, \tag{2.41}$$

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

Multiplying (2.40) by d_s and summing the results, we obtain

$$\sum_{i,j=1}^{m} \left(\sum_{s=2}^{m} (f_i(x_s^*) f_j(x_s^*))' d_s \right) d_{ij} = 0.$$

Substituting (2.40) in the above equation, we obtain

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

Since $(M(\xi_{\bar{\tau}}, z))^{-1}$ is a positive 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, (2.40) holds for $(i, j) = (m, m)$.

Define a vector ν by the equality

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

Certainly, $\nu^T f(x_i^*) = 0$, $i = 1, \ldots, m-1$, and we obtain from (2.40) that

$$\sum_{s=2}^{m} \nu^T f'(x_s^*) f_j(x_s^*) d_s = 0, \;\; j = 1, \ldots, m.$$

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

$$L_{(t)} \alpha = 0, \;\; t = 1, \ldots, m, \tag{2.42}$$

where $\alpha = \left(d_s \nu^T f'(x_s^*)\right)_{s=1}^{m-1}$; $L_{(t)}$ is obtained from the matrix $\left(f_i(x_j^*)\right)_{i,j=1}^{m,m-1}$ by rejecting the t-th line. It follows from (2.45) that $\det L_{(t)} = 0$, $t = 1, \ldots, m$, and it implies $\det \left(f_i(x_j^*)\right)_{i,j=1}^{m} = 0$. However, the last equality is impossible. ∎

Note that if $f_1(x), \ldots, f_{m-1}(x)$ generate a Chebyshev system on [a,b], then we need not use (2.41) and the points $x^*i, i = 2, \ldots, m$ need not to be support points of a locally D-optimal design in order for the matrix B be of full rank. This remark will be needed in the following for the consideration of MMEMS designs.

2.6.3 Proof of Theorem 2.2.3

Let us begin with the proof of part (I). A direct calculation shows that the matrix $J = J(\delta)$ is of the form

$$J = \begin{pmatrix} A & l \\ l^T & 0 \end{pmatrix},$$

where

$$\begin{aligned} A &= \left(\frac{\partial^2}{\partial\tau_i \partial\tau_j} \Phi(u,\delta) \right)_{i,j=1}^{m-1} \Bigg|_{u=\hat{u}(\delta)}, \\ l &= (l_1, \dots, l_{m-1})^T, \\ l_i &= \frac{R_1(\hat{\tau}_i)}{1-\delta} - \frac{R_2(\hat{\tau}_i)}{1+\delta},\ i = 1, \dots, m-1, \\ R_s(\hat{\tau}_i) &= \frac{\partial}{\partial\tau_i} \left(\det M(\xi_{\hat{\tau}}, \Lambda_{(s)}) \right)^{1/m} \\ &= \left(\det M(\xi_{\hat{\tau}}, \Lambda_{(s)}) \right)^{1/m} \\ &\quad f^T(\hat{\tau}, \Lambda_{(s)}) M^{-1}(\xi_{\hat{\tau}}, \Lambda_{(s)}) f(\hat{\tau}, \Lambda_{(s)}, \\ s &= 1,2,\ \Lambda_{(1)} = (1-\delta)c,\ \Lambda_{(2)} = (1+\delta)c. \end{aligned}$$

Let us prove that $l \neq (0, \dots, 0)^T$. Suppose, oppositely, that $l = (0, \dots, 0)^T$.

With $u = \hat{u}$, from the definition of $\hat{u}$, we have

$$\frac{\partial}{\partial\alpha} \Phi(u,\delta) = 0,\ \frac{\partial}{\partial\tau_i} \Phi(u,\delta) = 0,\ i = 1, \dots, m-1$$

for $u = \hat{u}$. From the first equality we obtain

$$\frac{\det M(\xi_{\hat{\tau}}, \Lambda_{(1)})}{1-\delta} = \frac{det M(\xi_{\hat{\tau}}, \Lambda_{(2)})}{1+\delta}.$$

Due to other $m-1$ equalities, we have

$$\frac{1}{1+\delta} R_2(\hat{\tau}_i) + \alpha \left\{ \frac{R_1(\hat{\tau}_i)}{1-\delta} - \frac{R_2(\hat{\tau}_i)}{1+\delta} \right\} = 0,\ i = 1, \dots, m-1.$$

Now, it follows from the supposition $l = (0, \dots, 0)^T$ that

$$\frac{\partial}{\partial\tau_i} \varphi(\tau, \Lambda_{(s)}) = 0,\ i = 1, \dots, m,\ s = 1,2$$

with $\tau = \hat{\tau}$.

A direct calculation shows that

$$\varphi(\tau, \Lambda_{(2)}) = \varphi\left(\frac{\tau}{1+\delta}(1-\delta), \Lambda_{(1)} \right) \frac{1+\delta}{1+\delta}.$$

Therefore,

$$\begin{aligned}&\frac{\partial}{\partial\tau_i}\varphi(\hat{\tau},\Lambda_{(1)})=0,\ i=1,\dots,m-1,\\&\frac{\partial}{\partial\tau_i}\varphi(h\hat{\tau},\Lambda_{(1)})=0,\ h=\frac{1+\delta}{1-\delta},\ i=1,\dots,m-1.\end{aligned}\tag{2.43}$$

However, in Melas (1978) it was proved that the equation system (2.43) has a unique solution in the set V. The contradiction obtained proves that $l\neq(0,\dots,0)^T$.

Let us now study the matrix A. Similar to the proof of Theorem 2.4.2, it can be proved that the matrix A has the form

$$A=E-\alpha B_{(1)}^T\mathcal{D}_{(1)}^{-1}B_{(1)}-(1-\alpha)B_{(2)}^T\mathcal{D}_{(2)}^{-1}B_{(2)},$$

where E is a diagonal matrix, $\mathcal{D}_{(1)}$ and $\mathcal{D}_{(2)}$ are positive definite, and $\alpha=\hat{\alpha}$.

Repeating the arguments from the proof of Theorem 2.3.1, obtain that the matrices $B_{(1)}$ and $B_{(2)})$ have full rank and $(E)_{ii}\leq 0$, $i=1,\dots,m-1$. Therefore the matrix A is negative definite and invertible.

Now, we have

$$\det J=-l^TAl\neq 0.$$

Now, assertion (I) of Theorem 2.2.3 follows from Theorem 2.4.1.

Let us prove part (II). From the general equivalence theorem for maximin efficient designs (see Dette, Haines and Imhof (2003) or Müller and Pazman (1998)) it follows that the MMEMS design $\xi_{\hat{\tau}}$ is MME design among all approximate designs if and only if the two following conditions are satisfied:

$$\begin{aligned}&\hat{\alpha}f^T(x,\Lambda_{(1)})M^{-1}(\xi_{\hat{\tau}},\Lambda_{(1)})f(x,\Lambda_{(1)})\\&\qquad+(1-\hat{\alpha})F^T(x,\Lambda_{(2)})M^{-1}(\xi_{\hat{\tau}},\Lambda_{(2)})f(x,\Lambda_{(2)})\leq m\end{aligned}\tag{2.44}$$

with $x\geq 0$, where $\Lambda_{(1)}=(1-\delta)c$, $\Lambda_{(2)}=(1+\delta)c$, and $\hat{\tau}=\hat{\tau}(\delta)$, and

$$\min_{\Lambda\in\Omega(\delta)}(\tau,\Lambda)=\min_{0\leq\alpha\leq 1}\alpha Q(\tau,\Lambda_{(1)})+(1-\alpha)Q(\tau,\Lambda_{(2)}),\tag{2.45}$$

where

$$Q(\tau,\Lambda)=\varphi(\tau,\Lambda)/\varphi(\tau^*(\Lambda),\Lambda).$$

In a vicinity of $\Lambda=c$ let LDMS designs be locally D-optimal among all approximate designs. Then, due to the standard continuity arguments, inequality (2.44) holds for sufficiently small δ.

In order to prove (2.45), we will need the following auxiliary result.

Lemma 2.6.2 *Consider a general function $\varphi:V\times\Omega\to R$, where $V\subset R^s$ and $\Omega\subset R^k$ are open sets.*

Suppose that the following assumptions are satisfied:

(a1) The function φ is positive and twice continuously differentiable.

(a2) For any $\Lambda \in \Omega$, the equation

$$g(\tau, \Lambda) = 0,$$

where $g(\tau,\Lambda) = \left(\frac{\partial}{\partial \tau_1}\varphi(\tau,\Lambda), \ldots, \frac{\partial}{\partial \tau_3}\varphi(\tau,\Lambda)\right)^T$ possess a unique solution $\tau^* = \tau^*(\Lambda)$.

(a3) For all $\Lambda \in \Omega$, the matrix

$$K = B^T J(\Lambda) B,$$

where

$$J(\Lambda) = \left(\tfrac{\partial^2}{\partial \tau_i \partial \tau_j}\varphi(\tau,\Lambda)\right)_{i,j=1}^{s}\Bigg|_{tau=\tau^*(\Lambda)},$$

$$B = \left(\tfrac{\partial \tau_i(\Lambda)}{\partial \lambda_j}\right)_{i,j=1}^{sk}, \ \lambda_j = (\Lambda)_j,$$

consists of negative elements.

Let $\Omega = \Omega(\delta)$, where

$$\Omega(\delta) = \left\{\Lambda \middle| \Lambda = (\lambda_1, \ldots, \lambda_k), (1-\delta)c_i \le \lambda_i \le (1+\delta)c_i, \ i = 1, \ldots, k\right\},$$

$0 < \delta < 1$, and

$$Q(\tau,\Lambda) = \varphi(\tau,\Lambda)/\varphi(\tau^*(\Lambda),\Lambda).$$

Then for τ sufficiently close to $\tau^*(\Lambda)$

$$\min_{\Lambda \in \Omega(\delta)} Q(\tau,\Lambda) = \min_{0 \le \alpha \le 1} \alpha(Q(\tau,\Lambda_{(1)}) + (1-\alpha)Q(\tau,\Lambda_{(2)}),$$

where $\Lambda_{(1)} = (1-\delta)c$, $\Lambda_{(2)} = (1+\delta)c$.

Proof of the lemma. The proof is similar to that of Proposition A1 in Dette, Melas and Pepelyshev (2003). A direct calculation shows that

$$\begin{aligned}\frac{\partial}{\partial \lambda_i}\varphi(\tau^*(\Lambda),\Lambda) &= \frac{\partial}{\partial \lambda_i}\varphi(\tau,\Lambda)\Big|_{\tau=\tau^*(\Lambda)} \\ &\quad + \sum_{j=1}^{s}\frac{\partial}{\partial \tau_j}\varphi(\tau,\Lambda)\Big|_{\tau=\tau^*(\Lambda)}\frac{\partial}{\partial \lambda_i}(\tau_j^*(\Lambda))\end{aligned}, \tag{2.46}$$

$i = 1, 2, \ldots, k$.

Due to assumption (a2), we obtain that

$$\begin{aligned}\frac{\partial^2}{\partial \lambda_i \partial \lambda_j}\varphi(\tau^*(\Lambda),\Lambda) &= \frac{\partial^2}{\partial \lambda_i \partial \lambda_j}\varphi(\tau,\lambda)\Big|_{tau=\tau^*(\Lambda)} \\ &\quad + \sum_{u,v=1}^{s}\frac{\partial^2}{\partial \tau_u \partial \tau_v}\varphi(\tau,\Lambda)\Big|_{rau=tau^*(\Lambda}\frac{\partial}{\partial \lambda_i}\tau_u^*(\Lambda)\frac{\partial}{\partial \lambda_j}\tau_v^*(\Lambda),\end{aligned} \tag{2.47}$$

$i, j = 1, \ldots, k$.

Now, it is easy to calculate

$$\frac{\partial^2}{\partial \lambda_i \partial \lambda_j} Q(\tau, \Lambda) = Q_{1ij}(\tau, \Lambda) + Q_{2ij}(\tau, \Lambda),$$

where $Q_{1ij}(\tau, \Lambda)$ is such that $Q_{1ij}(\tau^*(\Lambda), \Lambda) = 0$,

$$Q_{2ij}(\tau, \Lambda) = \left[\left(\frac{\partial^2}{\partial \lambda_i \partial \lambda_j} \varphi(\tau, \Lambda) \right) H(\Lambda) - \varphi(\tau, \Lambda) \frac{\partial^2}{\partial \lambda_i \partial \lambda_j} H(\Lambda) \right] \Big/ H^2(\Lambda),$$
$$H(\lambda) = \varphi(\tau^*(\Lambda), \Lambda),$$

$i, j = 1, 2, \ldots, k$.

From (2.47) and the above formulas it follows that the matrix

$$\left(\frac{\partial^2}{\partial \lambda_1 \partial \lambda_j} Q(\tau, \Lambda) \right)_{i,j=1}^{k} \Bigg|_{\tau = \tau^*(\Lambda)}$$

is equal to K.

Since all elements of this matrix are negative by assumption (a3), the minimum of $Q(\tau, \Lambda)$ by $\Lambda \in \Omega(\delta)$ is achieved at the set $\{\Lambda_{(1)}, \Lambda_{(2)}\}$ for sufficiently small δ and with τ sufficiently close to $\tau^*(\Lambda)$. This is equivalent to the assertion of the lemma. ∎

Now, let

$$\varphi(\tau, \Lambda) = \left(\frac{\det M(\xi_\tau, \Lambda)}{\prod_{i<j} (\lambda_i - \lambda_j)^8} \right)^{1/m}. \tag{2.48}$$

We can assume that the function in the points Λ with $\lambda_i = \lambda_j$ for some $i \neq j$ is codetermined with preserving the continuity (it can be done due to the discussion in Section 2.2).

Condition (a1) is evidently satisfied for this function and conditions (a2) and (a3) are proved in Melas (1978).

Thus, due to Lemma 2.6.2 for the function $\varphi(\tau, \Lambda)$ determined by (2.48) condition (2.45) takes place for sufficiently small δ. This completes the proof of part (II) of Theorem 2.2.3.

http://www.springer.com/978-0-387-98741-5

Functional Approach to Optimal Experimental Design
Melas, V.
2006, X, 338 p., Softcover
ISBN: 978-0-387-98741-5