

E -optimal designs for second order response surface models

Holger Dette

Ruhr-Universität Bochum

Fakultät für Mathematik

44780 Bochum, Germany

email: holger.dette@ruhr-uni-bochum.de

Yuri Grigoriev

St.-Petersburg State Electrotechnical University

Dept. of Computer Science

197376 St. Petersburg, Russia

email: yuri_grigoriev@mail.ru

March 12, 2014

Abstract

E -optimal experimental designs for a second order response surface model with $k \geq 1$ predictors are investigated. If the design space is the k -dimensional unit cube, Galil and Kiefer (1977a) determined optimal designs in a restricted class of designs (defined by the multiplicity of the minimal eigenvalue) and stated their universal optimality as a conjecture. In this paper we prove this claim and show that these designs are in fact E -optimal in the class of all approximate designs. Moreover, if the design space is the unit ball, E -optimal designs have not been found so far and we also provide a complete solution to this optimal design problem.

The main difficulty in the construction of E -optimal designs for the second order response surface model consists in the fact that for the multiplicity of the minimum eigenvalue of the “optimal information matrix” is larger than one (in contrast to the case $k = 1$) and as a consequence the corresponding optimality criterion is not differentiable at the optimal solution. These difficulties are solved by considering nonlinear Chebyshev approximation problems, which arise from a corresponding equivalence theorem. The extremal polynomials which solve these Chebyshev problems are constructed explicitly leading to a complete solution of the corresponding E -optimal design problems.

Keywords and phrases: response surface models, optimal designs, E -optimality, extremal polynomial, duality, nonlinear Chebyshev approximation

AMS Subject classifications: Primary 62K05; Secondary 41A50

1 Introduction

Response surface methodology has become a standard tool in the analysis of experimental data. These models are used to study the influence of several input factors on a response variable by approximating complex functional relationships by “simple” linear or quadratic multivariate polynomial regression models, which are usually denoted as first or second order response surface models [see for example Myers and Montgomery (2002)]. Numerous authors have worked on the construction of efficient and optimal experimental designs for response surface models. For first-order models 2^k factorial and fractional factorial 2^{k-p} designs of resolution III are optimal with respect to the D -, G - and I -optimality criteria [see Anderson-Cook et al. (2009)]. On the other hand, for the second order response surface model the situation is more complicated and intuitively reasonable designs with a “simple” structure such as central composite designs are not optimal.

For this model approximate designs in the sense of Kiefer (1974) have been investigated by several authors, where the methodology and optimal designs differ by the design space and optimality criterion under consideration (typical a k -dimensional cube, ball or simplex). D -optimal approximate designs for the second order polynomial regression model on the ball and cube have been determined explicitly by Kiefer (1959, 1961), Kiefer and Wolfowitz (1959), Kono (1962), Farrel et al. (1967) [see also Rafa]lowicz and Myszka (1988), Lim and Studden (1988) and Dette and Röder (1997) who determined optimal product designs for multivariate polynomial regression models in more general situations]. In particular, it is shown that D -optimal designs on a ball are at the same time rotatable designs. Considerably less attention has been paid to other optimality criteria. Laptev (1974), Denisov and Popov (1976) and Golikova and Pantchenko (1977) investigated A - and Q -optimal designs numerically, Galil and Kiefer (1977b) determined numerically rotatable optimal designs for the second order response surface model, while Draper et al. (2000) and Draper and Pukelsheim (2003) investigated optimal design problems in second order mixture models. On the other hand, the explicit determination of optimal designs in the class of all approximate designs with respect to other criteria than the D -criterion seems to be a very hard problem, which has only been solved in rare circumstances.

In this paper we study E -optimal designs for the second order response surface models on the k -dimensional cube and ball. Among Kiefer’s Φ_p -criteria [see Kiefer (1974)] the E -optimality criterion is not differentiable if the multiplicity of the minimum eigenvalue of the information matrix of the optimal design is larger than 1. This property makes the determination of E -optimal designs to an extremely hard and challenging problem. In fact an analytical construction of E -optimal designs for linear regression models is very difficult and has only been

achieved in the one-dimensional case for a limited number of linear and nonlinear models [see Melas (1982), Dette (1993), Pukelsheim and Studden (1993), Dette and Haines (1994), among others]. For models with more than one predictor results can only be found sporadically in the literature. For example, Cheng (1987) and Dette and Studden (1993) identified E -optimal spring balance and chemical balance weighing designs. Galil and Kiefer (1977a) considered the second order response surface model on the cube with k predictors and determined the E -optimal designs in the class of all designs, for which the corresponding information matrix has a minimum eigenvalue of multiplicity $k(k+1)/2$. However, to our best knowledge, the answer to the question, if these designs are in fact E -optimal in the class of all designs is still open. For the ball the situation is even worse, and only E -optimal designs in the class of all rotatable designs are available [see for example Galil and Kiefer (1977b)]. These designs are in fact not globally optimal and the determination of E -optimal designs for the second order response surface model on the ball is an open and challenging problem.

The goal of the present paper is to provide complete answers to these questions and to characterize the structure and properties of E -optimal designs for the second order response surface model. Our approach relies on a specific duality result for E -optimal designs, which relates the optimal design problem to a nonlinear Chebyshev approximation problem [see Melas (1982, 2006) or Pukelsheim (2006)]. In the dual problem one has to determine a nonnegative polynomial with minimal sup-norm in a specific class of nonnegative (multivariate) polynomials, that is

$$(1.1) \quad \mathcal{P} = \{f^T(x)Zf(x) \mid \text{trace}(Z) = 1; Z \geq 0\},$$

where x denotes the k -dimensional predictor, $f(x)$ is the vector of regression functions in the second order response surface model and Z is a nonnegative definite matrix of appropriate dimension. This Chebyshev approximation problem is nonlinear and therefore extremely hard to solve explicitly. For the solution of the E -optimal design problem this “optimal” polynomial, which is called extremal polynomial throughout this paper, will be constructed explicitly in Section 3 and 4 if the design space is the cube and ball, respectively. As a consequence, we are able to provide a complete solution of these E -optimal design problems. In general there exist several E -optimal designs which usually have a large number of support points. For this reason particular attention is paid to the problem of constructing E -optimal designs with a small number of support points.

2 Optimal designs for response surface models

We consider the common linear regression model of the form

$$(2.1) \quad \mathbb{E}(Y|x) = f^T(x)\theta,$$

where Y denotes the (one-dimensional) response and the explanatory variable x varies in a compact design space, say $\mathcal{X} \subset \mathbb{R}^k$. In (2.1) the vector $f(x) = (f_1(x), \dots, f_m(x))^T \in \mathbb{R}^m$ is the vector of regression functions and $\theta = (\theta_1, \dots, \theta_m)^T \in \mathbb{R}^m$ denotes a vector of unknown parameters. We assume that N independent observations are available according to the model (2.1) where at each experimental condition x the response y is a realization of a normal distributed random variable Y with expectation given by (2.1) and (constant) variance $\sigma^2 > 0$. An approximate designs in the sense of Kiefer (1974) is defined as probability measure on the design space \mathcal{X} with finite support. The support points, say $x_{(1)}, \dots, x_{(s)}$, of an approximate design ξ define the locations where observations are taken, while the weights give the corresponding relative proportions of total observations to be taken at these points. If the design ξ has masses $\omega_i > 0$ at the different points $x_{(i)}$ ($i = 1, \dots, s$) and N observations can be made by the experimenter, the quantities $\omega_i N$ are rounded to integers, say N_i , satisfying $\sum_{i=1}^s N_i = N$, and the experimenter takes N_i observations at each location $x_{(i)}$ ($i = 1, \dots, s$). The information matrix of an approximate design ξ is defined by

$$(2.2) \quad M(\xi) = \int_{\mathcal{X}} f(x)f^T(x)d\xi(x) \in \mathbb{R}^{m \times m},$$

and it is well known [see Jennrich (1969)] that under appropriate assumptions of regularity (in particular $\det(M(\xi)) > 0$ and $\lim_{N_i, N \rightarrow \infty} N_i/N = \omega_i > 0$) the covariance matrix of the least squares estimator is approximately given by $\sigma^2 M^{-1}(\xi)/N$, where N denotes the total sample size.

Optimal designs maximize an appropriate statistical meaningful functional, say Φ , of the information matrix. Among the the numerous criteria which have been proposed in the literature for this purpose [see Silvey (1980), Pázman (1986) or Pukelsheim (2006) among others] we consider in this paper the E -optimality criterion

$$(2.3) \quad \Phi_{-\infty}(\xi) = \lambda_{\min}(M(\xi)).$$

This criterion arises as a special case of Kiefer's Φ_p -optimality criteria, which are defined for $p \in (-\infty, 1]$ as

$$(2.4) \quad \Phi_p(M) = [m^{-1}\text{tr}(M^p(\xi))]^{1/p} = \left(m^{-1} \sum_{i=1}^m \lambda_i^p(M(\xi)) \right)^{1/p},$$

that is $\Phi_{-\infty}(\xi) = \lim_{p \rightarrow -\infty} \Phi_p(\xi)$ [see Kiefer (1974)]. In equation (2.4) the quantities $\lambda_1(M(\xi)), \dots, \lambda_m(M(\xi))$ denote the eigenvalues of the information matrix $M(\xi)$ and $\lambda_{\min}(M(\xi))$ its corresponding minimum eigenvalue. In contrast to the Φ_p -criteria with $p \in (-\infty, 1]$ the E -optimality criterion is not differentiable if the multiplicity of the minimum eigenvalue of the matrix $M(\xi)$ is larger than 1 and this property makes the determination of E -optimal designs to an extremely hard problem. In fact, E -optimal designs have been determined for a limited number of linear and nonlinear regression models [see the references cited in the introduction]. An important tool for the determination of E -optimal designs is the following equivalence theorem which has been proved by several authors [see Melas (1982) or Pukelsheim (2006) for example].

Theorem 2.1 *Let ξ^* denote a design and $\lambda_{\min}(M(\xi^*))$ the minimum eigenvalue of the information matrix $M(\xi^*)$ with multiplicity s . The design ξ^* is E -optimal if and only if there exist orthonormal eigenvectors q_0, \dots, q_{s-1} of the matrix $M(\xi^*)$ corresponding to $\lambda_{\min}(M(\xi^*))$ and nonnegative weights w_0, \dots, w_{s-1} with sum 1 such that the “extremal polynomial”*

$$d(x, \xi) = f^T(x)(q_0, \dots, q_{s-1}) \operatorname{diag}(w_0, \dots, w_{s-1})(q_0, \dots, q_{s-1})^T f(x) = \sum_{i=0}^{s-1} w_i (f^T(x)q_i)^2$$

satisfies for all $x \in \mathcal{X}$ the inequality

$$(2.5) \quad d(x, \xi) \leq \lambda_{\min}(M(\xi^*)).$$

Moreover, the maximum on the left hand side of (2.5) is attained at the support points of the E -optimal design ξ^* .

Remark 2.1 It follows from general equivalence theory developed in convex design theory [see Pukelsheim (2006)] that there exists a duality between the E -optimal design problem and a nonlinear Chebyshev approximation problem, that is

$$(2.6) \quad \max_{\xi} \lambda_{\min}(M(\xi)) = \min_{P_Z \in \mathcal{P}} \max_{x \in \mathcal{X}} |P_Z(x)|$$

where $\mathcal{P} = \{P_Z(x) = f^T(x)Zf(x) \mid Z \in \mathbb{R}^{m \times m}, Z \geq 0, \operatorname{trace}(Z) = 1\}$ denotes a subset of the nonnegative “polynomials”. In fact if there is equality in (2.6) for a pair (ξ^*, Z^*) , then ξ^* is an E -optimal design and P_{Z^*} a solution of the nonlinear Chebyshev approximation problem. This explains the name “extremal polynomial” in Theorem 2.1.

The second order response surface model with a k -dimensional predictor appears as a special case of model (2.1), that is

$$(2.7) \quad \mathbb{E}[Y|x] = \sum_{\|\alpha\|_1=0}^2 \theta_{\alpha} x^{\alpha},$$

where $\alpha = (\alpha_1, \dots, \alpha_k)^T \in \{0, 1, 2\}^k$ is a multi-index $x^\alpha = x^{\alpha_1} \dots x^{\alpha_k}$ and $\|\alpha\|_1 = \alpha_1 + \dots + \alpha_k$. In this case the corresponding vector of regression function in the general linear model (2.1) is given by

$$(2.8) \quad f(x) = (1, x_1^2, \dots, x_k^2, x_1, \dots, x_k, x_1x_2, \dots, x_{k-1}x_k) \in \mathbb{R}^m,$$

where $m = \frac{(k+1)(k+2)}{2}$, $x = (x_1, \dots, x_k)^T$. In the following section we consider optimal designs for the second order regression model (2.7), where the design spaces are the unit ball with respect to the maximum norm $\|\cdot\|_\infty$ and the Euclidean norm $\|\cdot\|_2$, that is

$$(2.9) \quad \mathcal{X} = \mathcal{B}_\infty(1) := \{x \in \mathbb{R}^k \mid \|x\|_\infty \leq 1\}, \quad \mathcal{X} = \mathcal{B}_2(1) := \{x \in \mathbb{R}^k \mid \|x\|_2 \leq 1\}.$$

It turns out that designs with certain symmetry properties play a particular role for the construction of E -optimal designs. Throughout this paper we call a design *symmetric* if for any $(\alpha_1, \dots, \alpha_k) \in \{0, 1, 2\}^k$ with $\|\alpha\|_1 = |\alpha_1| + \dots + |\alpha_k| \leq 2$ the moments

$$\int_{\mathcal{X}} x_1^{\alpha_1}, \dots, x_k^{\alpha_k} \xi(dx)$$

are invariant with respect to all permutations of $\alpha_1, \dots, \alpha_k$ and vanish if there is at least one odd index among $\alpha_1, \dots, \alpha_k$. In the following discussion let $I_\ell \in \mathbb{R}^{\ell \times \ell}$ denote the identity matrix and $1_\ell = (1, \dots, 1)^T \in \mathbb{R}^\ell$ denotes the vector with all elements equal to 1, then a straightforward calculation shows that the information matrix of a symmetric design in model (2.7) is of the form

$$(2.10) \quad M(\xi) = \int_{\mathcal{X}} f(x) f^T(x) \xi(dx) = \begin{pmatrix} 1 & a1_k^T & 0 & 0 \\ a1_k & H & 0 & 0 \\ 0 & 0 & aI_k & 0 \\ 0 & 0 & 0 & bI_{\frac{k(k-1)}{2}} \end{pmatrix} \in \mathbb{R}^{m \times m},$$

where $m = \frac{(k+1)(k+2)}{2}$, $H = H(c; b) = (c - b)I_k + b1_k1_k^T \in \mathbb{R}^{k \times k}$ denotes a circulant matrix with diagonal and off-diagonal elements c and b , respectively, and the entries a, b and c in (2.10) are given by

$$(2.11) \quad a = \int_{\mathcal{X}} x_1^2 \xi(dx), \quad b = \int_{\mathcal{X}} x_1^2 x_2^2 \xi(dx), \quad c = \int_{\mathcal{X}} x_1^4 \xi(dx).$$

It is easy to see that the entries in this matrix satisfy the inequalities

$$(2.12) \quad 1 > a \geq c > b > 0, \quad c + b(k - 1) > ka^2.$$

Designs with information matrix of the form (2.10) will serve as candidates for E -optimal designs. Consider for example the case $k = 1$, where model (2.7) reduces to the well known one-dimensional quadratic regression model $\theta_0 + \theta_1 x^2 + \theta_2 x$. If the design space is given by $\mathcal{X} = [-1, 1]$ and the design ξ puts masses $1/5$, $1/5$ and $3/5$ at the points -1 , 1 and 0 , respectively, the corresponding information matrix is given by

$$M(\xi^*) = \begin{pmatrix} 1 & \frac{2}{5} & 0 \\ \frac{2}{5} & \frac{2}{5} & 0 \\ 0 & 0 & \frac{2}{5} \end{pmatrix}.$$

It was shown by Kiefer (1974) that this design is in fact E -optimal for the univariate quadratic regression model and the minimum eigenvalue $\lambda_{\min} = \frac{1}{5}$ has multiplicity $s = 1$. For a similar statement in the univariate polynomial regression model of degree $d \geq 2$ see Pukelsheim and Studden (1993).

However, in the case $k \geq 2$ the multiplicity of the minimum eigenvalue of the matrix (2.10) is larger than 1 and as consequence the corresponding optimality criterion is not differentiable at the matrix $M(\xi)$ given by (2.10). This makes the determination of E -optimal designs substantially more difficult. For example, Galil and Kiefer (1977a) determined the E -optimal design on the cube $\mathcal{B}_{\infty}(1)$ in the subclass of all designs with information matrix of the form (2.10), where its minimum eigenvalue has multiplicity $\frac{k(k+1)}{2}$ (these calculations will be briefly presented at the beginning of the following section). To our best knowledge the question, if the solution obtained by these authors in the restricted class yields in fact an E -optimal design for the second order response surface model in the class of all approximate designs on the cube, has not been answered. Moreover, the E -optimal design problem for second order regression models seems to be completely unsolved if the design space is given by the unit ball $\mathcal{B}_2(1)$.

In the following two sections we will present a complete solution to these problems. For this purpose we proceed in the following sections in two steps:

- (I) In a first step a candidate for the E -optimal design in the class of all designs with information matrix of the form (2.10) is identified. If the design space is given by the cube our arguments coincide with those of Galil and Kiefer (1977a) and are presented here for the sake of completeness.
- (II) In a second step the E -optimality of the candidate design found by Galil and Kiefer (1977a) is proved by an application of Theorem 2.1. This requires the determination of an appropriate basis of the eigenspace corresponding to the minimum eigenvalue of $M(\xi)$ and the construction of the corresponding extremal polynomial in (2.6).

The E -optimal designs for the second order response surface model will be identified in terms of the masses that they assign to specific sets which depend on the design space under consideration. Because in many applications it is desirable to obtain optimal designs with a minimal number of support points we add a third step if the design space is the cube, that is

(III) Identification of designs with a minimal number of support points.

3 E -optimal designs on the cube

In this section we consider the second order response surface model (2.1) on the design space $\mathcal{X} = \mathcal{B}_\infty(1) = [-1, 1]^k$. We start with a determination of a “good” candidate for an E -optimal symmetric design. Our arguments are similar to those given in Galil and Kiefer (1977a) and presented here for the sake of completeness (note that these authors only identified the candidate design and in the following we will prove its optimality in the class of all approximate designs). Observing the representation of the corresponding information matrix (2.10) the eigenvalues of the matrix $M(\xi)$ are given by a , b , and the eigenvalues by its upper $(k + 1) \times (k + 1)$ block,

$$(3.1) \quad M_{11}(\xi) = \begin{pmatrix} 1_k & a1_k^T \\ a1_k & H \end{pmatrix},$$

where $H = H(c; b) = (c - b)I_k + b1_k1_k^T$. Define $D = [1 - c - (k - 1)b]^2 + 4ka^2 > 0$, then all eigenvalues of the information matrix of a symmetric E -optimal design are given by

$$(3.2) \quad \lambda_0 = \frac{1 + c + (k - 1)b + \sqrt{D}}{2}, \quad \lambda_1 = \frac{1 + c + (k - 1)b - \sqrt{D}}{2}, \\ \lambda_2 = \dots = \lambda_k = c - b, \quad \lambda_{k+1} = \dots = \lambda_{2k} = a, \quad \lambda_{2k+1} = \dots = \lambda_m = b.$$

Note that $\lambda_0 > \lambda_1$ and that λ_1 and λ_2 are increasing functions of c . Observing (2.12) we therefore obtain $c = a$ and the problem of maximizing the minimum eigenvalue of $M(\xi)$ reduces to the maximization of

$$(3.3) \quad \lambda_{\min}(M(\xi)) = \min \left\{ \frac{1 + a + (k - 1)b - \sqrt{D}}{2}, a - b, a, b \right\} = \min \left\{ \frac{1 + a + (k - 1)b - \sqrt{D}}{2}, a - b \right\},$$

where the constant D is now represented as $D = [1 - a - (k - 1)b]^2 + 4ka^2$ and the second equality in (3.3) follows from $0 < a - b < a$ [see (2.12)]. Motivated by the solution of similar maximin problems, we suppose that $\lambda_1 = \lambda_2$, which gives

$$\lambda_1 = \lambda_2 = a - \frac{-(1 - a) + \sqrt{(1 - a)^2 + 4a^2}}{2} = \frac{1 + a - \sqrt{(1 - a)^2 + 4a^2}}{2}.$$

Maximization of (3.3) subject to $\lambda_1 = \lambda_2$ finally yields for the elements of the matrix $M(\xi)$ $a = \frac{2}{5}$, $b = \frac{1}{5}$, and for its eigenvalues

$$(3.4) \quad \lambda_0 = 1 + \frac{k}{5}, \quad \lambda_1 = \dots = \lambda_k = \frac{1}{5}, \quad \lambda_{2k+1} = \dots = \lambda_{\frac{k(k-1)}{2}} = \frac{1}{5}, \quad \lambda_{k+1} = \dots = \lambda_{2k} = \frac{2}{5},$$

where the corresponding multiplicities of $\lambda_0, \lambda_1, \lambda_{k+1}$ are given by $1, \frac{k(k+1)}{2}$ and k , respectively. Hence we obtain as a candidate for an E -optimal information matrix the matrix $M(\xi^*)$ in (2.10) with $a = c = \frac{2}{5}$, $b = \frac{1}{5}$, where the minimum eigenvalue is given by $\lambda_{\min}(M(\xi^*)) = \frac{1}{5}$. This means that the information matrix under consideration has a minimal eigenvalue with multiplicity $\frac{k(k+1)}{2} \geq 3$ whenever $k \geq 2$. The following result gives an answer to the question if the determined values for a and b yield in fact to an E -optimal information matrix.

Theorem 3.1 *Any design ξ^* with an information matrix $M(\xi^*)$ of the form (2.10) and $a = c = \frac{2}{5}$ $b = \frac{1}{5}$ is E -optimal for the second order response surface model (2.7) on the k -dimensional unit cube. In particular Theorem 2.1 holds with*

$$(3.5) \quad d(x, \varepsilon) = \frac{1}{5} \left(1 - \frac{4}{k} \sum_{i=1}^k x_i^2 (1 - x_i^2) \right).$$

The Proof of Theorem 3.1 is complicated and deferred to Section 5.1. Note that in contrast to the D -optimality criterion the optimal values for a and b do not depend on a dimension of the design space. This fact has been independently observed by Denisov and Popov (1976) and Galil and Kiefer (1977a), who identified the correct E -optimal information matrix but did not prove its optimality.

In the next step we determine designs with corresponding information matrix specified in Theorem 3.1. For this purpose we call a point $x \in \mathbb{R}^k$ a barycenter of depth $0 \leq j \leq k$ if j coordinates are equal to 0 and the remaining $k - j$ coordinates are equal to ± 1 [see Galil and Kiefer (1977a)]. The set of all barycenters of depth r is denoted E_r and for its cardinality we introduce the symbol

$$(3.6) \quad n_r := |E_r| = \binom{k}{r} 2^{k-r}, \quad r = 0, 1, \dots, k.$$

It was shown by Kiefer (1960) and Farrel et al. (1967) that the support of every Φ_p -optimal design for the second order response surface model on the cube is a subset of the set

$$(3.7) \quad E = \bigcup_{j=0}^k E_j.$$

Moreover, there always exists a symmetric optimal design. Throughout this section we will describe these symmetric designs on the cube in terms of the $(k + 1)$ -dimensional vector $\xi =$

(ξ_0, \dots, ξ_k) , where ξ_i represents the mass assigned by the design to the set E_i of barycenters of depth i , that is $\xi_i = \xi(E_i)$ ($i = 0, \dots, k$). It turns out that there always exists an E -optimal design supported at at most three sets E_i . For this purpose we define for integers $0 \leq r_1 < r_2 < r_3 \leq k$ the matrix

$$A_{r_1, r_2, r_3} = \begin{pmatrix} 1 & 1 & 1 \\ \frac{k-r_1}{k} & \frac{k-r_2}{k} & \frac{k-r_3}{k}, \\ \frac{k-r_1}{k} \frac{k-r_1-1}{k-1} & \frac{k-r_2}{k} \frac{k-r_2-1}{k-1} & \frac{k-r_3}{k} \frac{k-r_3-1}{k-1} \end{pmatrix}$$

Lemma 3.1 *There exists integers $0 \leq r_1 < r_2 < r_3 \leq k$ such that the system of linear equations*

$$(3.8) \quad A_{r_1, r_2, r_3} \xi = \left(1, \frac{2}{5}, \frac{1}{5}\right)^T$$

has a unique solution $\xi^* = (\xi_1^*, \xi_2^*, \xi_3^*)^T$ satisfying $\xi_i^* \geq 0$, $\sum_{i=1}^3 \xi_i = 1$. Any design with masses

$$(3.9) \quad \xi(E_{r_i}) = \xi_i^*, \quad i = 1, 2, 3,$$

is E -optimal for the second order response surface model (2.7).

Proof. Let ξ denote a symmetric design and note that the moments in the matrix $M(\xi)$ defined in (2.10) have the representation

$$(3.10) \quad 1 = \sum_{r=0}^k \xi_r, \quad a = \sum_{r=0}^{k-1} a_r \xi_r, \quad b = \sum_{r=0}^{k-2} b_r \xi_r,$$

where $\xi_r = \xi(E_r)$ is the measure of the set E_r of barycenters of depth r and

$$(3.11) \quad a_r := \binom{k-1}{r} 2^{k-r}, \quad r \in \{0, \dots, k-1\}, \quad b_r := \binom{k-2}{r} 2^{k-r}, \quad r \in \{0, \dots, k-2\}.$$

By (3.10) and a remark on page 124 of Galil and Kiefer (1977a) there exists symmetric design ξ and three sets E_{r_1}, E_{r_2} and E_{r_3} such that (3.10) is satisfied for $a = \frac{2}{5}$ and $b = \frac{1}{5}$. A simple calculation shows that in this case the system of equations in (3.10) is equivalent to (3.8), which has a unique solution because $\det(A) = \frac{(r_1-r_2)(r_1-r_3)(r_2-r_3)}{k^2(k-1)} \neq 0$. \square

It should be noted that not any solution of (3.8) will yield a vector of admissible weights $(\xi_{r_1}, \xi_{r_2}, \xi_{r_3}) = (\xi(E_{r_1}), \xi(E_{r_2}), \xi(E_{r_3}))$ (some components could be negative). Moreover, in general there exist many triples (r_1, r_2, r_3) , such that the system (3.8) has a solution with non-negative components and any such triple yields to at least one symmetric E -optimal design.

For example, if (r_1, r_2, r_3) is such a triple with corresponding solution $(\xi(E_{r_1}), \xi(E_{r_2}), \xi(E_{r_3}))$ of (3.8), then a design ξ which assigns masses

$$\omega_{r_i, j} = \xi(\{x_{r_i, j}\}) = \frac{\xi(E_{r_i})}{n_{r_i}}; \quad j = 1, \dots, n_{r_i}; \quad i = 1, 2, 3;$$

to all points $x_{(r_i, 1)} \dots x_{(r_i, n_{r_i})} \in E_{r_i}$ is an E -optimal design for the second order response surface model (2.7) on the unit cube $[-1, 1]^k$, where $n_j = \binom{k}{j} 2^{k-j}$ denotes the number of elements of the set E_j ($j = 0, \dots, k$). The number of support points of such a design is given by

$$N(r_1, r_2, r_3) = \sum_{i=1}^3 \binom{k}{r_i} 2^{k-r_i}$$

and usually rather large. For this reason it is of interest to find designs with a minimal number of support points [see Farrel et al. (1967) or Pesotchinsky (1975)]. A reasonable approach to this problem is to look for E -optimal designs which are supported at only **two** sets of barycenters, say E_{r_1} and E_{r_2} . Because it can easily be shown that for a triple (r_1, r_2, r_3) with an admissible solution of (3.8) the weights $\xi(E_{r_i})$ are given by

$$(3.12) \quad \xi(E_{r_1}) = \frac{1}{5} \cdot \frac{2k^2 + k - 3k(r_2 + r_3) + 5r_2 r_3}{(r_2 - r_1)(r_3 - r_1)}, \quad i = 1, 2, 3,$$

it follows that symmetric E -optimal designs supported at only two sets of barycenters can be obtained from the *Diophantine equations*

$$(3.13) \quad 2k^2 + k - 3k(s + t) + 5st = 0$$

for $s, t = 0, \dots, k$. These equations have been solved numerically by Galil and Kiefer (1977a) if $k \leq 25$ (see Table 1 in this reference). It should be pointed here that there does not always exist a solution of (3.13) (for example for $k = 2, 6$, or 8). Moreover, in general it is not clear that a solution of (3.13) necessarily yields to an E -optimal design with a minimal number of support points. For this reason we display in Table 1 the E -optimal symmetric designs with a minimal number of support points for second order response surface models with $k \leq 24$ predictors. For example, if $k = 5$, the design with a minimal number of support points in only two sets has $N(2, 5) = 81$ support points in the set E_2 and E_5 [see Galil and Kiefer (1977a)], while the design with the minimal number of $N(0, 3, 5) = 73$ support points in the sets E_0 , E_3 and E_5 .

Remark 3.2 Based on our numerical results we found a remarkable structure for the E -optimal designs with a minimal number of support points for the second order response surface model with k predictors, whenever $k \neq 3$. The E -optimal design for the second order response surface

k	(r, s, t)	$\xi(E_r)$	$\xi(E_s)$	$\xi(E_t)$	k	(r, s, t)	$\xi(E_r)$	$\xi(E_s)$	$\xi(E_t)$
1	$(0, 1, -)$	$\frac{2}{5}$	$\frac{3}{5}$	—	13	$(0, 9, -)$	$\frac{2}{15}$	$\frac{13}{15}$	—
2	$(0, 1, 2)$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$	14	$(0, 9, 14)$	$\frac{25}{225}$	$\frac{182}{225}$	$\frac{18}{225}$
3	$(-, 1, 3)$	—	$\frac{3}{5}$	$\frac{2}{5}$	15	$(0, 10, 15)$	$\frac{3}{25}$	$\frac{21}{25}$	$\frac{1}{25}$
4	$(0, 3, -)$	$\frac{1}{5}$	$\frac{4}{5}$	—	16	$(0, 11, -)$	$\frac{7}{55}$	$\frac{48}{55}$	—
5	$(0, 3, 5)$	$\frac{2}{15}$	$\frac{10}{15}$	$\frac{3}{15}$	17	$(0, 11, 17)$	$\frac{18}{165}$	$\frac{136}{165}$	$\frac{11}{165}$
6	$(0, 4, 6)$	$\frac{3}{20}$	$\frac{15}{20}$	$\frac{2}{20}$	18	$(0, 12, 18)$	$\frac{7}{60}$	$\frac{51}{60}$	$\frac{2}{60}$
7	$(0, 5, -)$	$\frac{4}{25}$	$\frac{21}{25}$	—	19	$(0, 13, -)$	$\frac{8}{65}$	$\frac{57}{65}$	—
8	$(0, 5, 8)$	$\frac{9}{75}$	$\frac{56}{75}$	$\frac{10}{75}$	20	$(0, 13, 20)$	$\frac{49}{455}$	$\frac{380}{455}$	$\frac{26}{455}$
9	$(0, 6, 9)$	$\frac{2}{15}$	$\frac{12}{15}$	$\frac{1}{15}$	21	$(0, 14, 21)$	$\frac{4}{35}$	$\frac{30}{35}$	$\frac{1}{35}$
10	$(0, 7, -)$	$\frac{1}{7}$	$\frac{6}{7}$	—	22	$(0, 15, -)$	$\frac{3}{25}$	$\frac{22}{25}$	—
11	$(0, 7, 11)$	$\frac{8}{70}$	$\frac{55}{70}$	$\frac{7}{70}$	23	$(0, 15, 23)$	$\frac{32}{300}$	$\frac{253}{300}$	$\frac{15}{300}$
12	$(0, 8, 12)$	$\frac{5}{40}$	$\frac{33}{40}$	$\frac{2}{40}$	24	$(0, 16, 24)$	$\frac{9}{80}$	$\frac{69}{80}$	$\frac{2}{80}$

Table 1: *Symmetric E-optimal designs with a minimal number of support points for second order response surface models with $k \leq 24$ predictors.*

$l = +1$			$l = 0$			$l = -1$		
$\xi(E_0)$	$\xi(E_s)$	$\xi(E_k)$	$\xi(E_0)$	$\xi(E_s)$	$\xi(E_k)$	$\xi(E_0)$	$\xi(E_s)$	$\xi(E_k)$
$\frac{1}{5} \cdot \frac{q+2}{2q+1}$	$\frac{3}{5} \cdot \frac{3q+1}{2q+1}$	0	$\frac{1}{5} \cdot \frac{q+3}{2q}$	$\frac{3}{5} \cdot \frac{3q-1}{2q}$	$\frac{1}{5q}$	$\frac{1}{5} \cdot \frac{q}{2q-1}$	$\frac{1}{5} \cdot \frac{(3q-1)(3q-2)}{q(2q-1)}$	$\frac{2}{5q}$

Table 2: Conjecture for the structure of E -optimal designs with a minimal number of support points for second order response surface models with $k = 1, 2$ and $k \geq 4$ predictors, where $k = 3q + l$ and $s = 2q + l$ and $l = 0, \pm 1$.

model with a minimal number of support points is always supported at the sets E_0 and E_k and a third set E_s . If $k = 3q + l$ where $l = 0, \pm 1$, then $s = 2q + l$. The particular structure is displayed in Table 2, which also contains the weights assigned by the E -optimal design to these sets.

Example 3.1 Galil and Kiefer (1977a) presented in Table 2 of their paper E -optimal designs (obtained as limits of Φ_p -optimal designs as $p \rightarrow -\infty$). Note that not all designs in this class have the minimal number of support points. For example, if $k = 6$ the E -optimal design obtained by Galil and Kiefer (1977a) puts masses 0.040, 0.400 and 0.560 at the sets E_0 , E_2 and E_5 respectively and has 316 support points. The E -optimal design obtained from Table 2 puts masses $\xi(E_0) = 0.15$, $\xi(E_4) = 0.75$, $\xi(E_6) = 0.10$ and has only 125 support points.

4 E -optimal designs on the unit ball

In this section we consider the E -optimal design problem for the second order response surface model on the k -dimensional ball $\mathcal{B}_2(1) = \{x \in \mathbb{R}^k : \|x\|_2 \leq 1\}$. The general strategy for the solution of the optimal design problem will be similar as the one given for the cube and we start identifying a good candidate for the E -optimal design. If the design space is the ball, then the sets E_{r_i} of barycenters of depth r_i will be replaced by three sets F_0 , F_{k-1} and F_k as candidate sets for the support of E -optimal designs. Here F_0 consists of the 2^k vertices $x = (\pm \frac{1}{\sqrt{k}}, \dots, \pm \frac{1}{\sqrt{k}})^T \in \mathbb{R}^k$ of the cube $\mathcal{B}_\infty(\sqrt{2})$ inscribed in k -dimensional ball $\mathcal{B}_2(1)$, F_{k-1} consists of the centers $\pm e_i$ of the $(k-1)$ -dimensional faces of $\mathcal{B}_\infty(1)$ (here $e_i = (0, \dots, 0, 1, 0, \dots, 0)^T$ denotes the i th unit vector) and F_k contains only the center of the ball. Note that the cardinality of these sets are given by

$$(4.1) \quad |F_0| = 2^k, \quad |F_{k-1}| = 2k, \quad |F_k| = 1.$$

As a consequence, there is no necessity to search for the minimally supported designs on the unit ball.

Consider a symmetric design ξ which is supported on the sets F_0 , F_{k-1} and F_k introduced in the previous paragraph. Its information matrix $M(\xi)$ in the second order response surface model (2.1) is of the form (2.10) with corresponding eigenvalues given by (3.2) where $D = [1 - (c - b) - kb]^2 + 4ka^2 > 0$. Moreover, from the definition of ξ we have for the entries defined in the matrix (2.10)

$$(4.2) \quad a = k^{-1}\xi(F_0) + \xi(F_{k-1}), \quad b = k^{-2}\xi(F_0), \quad c = k^{-2}\xi(F_0) + \xi(F_{k-1}),$$

and it now follows

$$(4.3) \quad \xi(F_{k-1}) = a - kb = c - b.$$

Substituting this identity into expression (3.2) for λ_1 yields

$$(4.4) \quad \lambda_1 = \frac{1 + a - \sqrt{(1 - a)^2 + 4ka^2}}{2}.$$

Therefore, the problem of determining an E -optimal (symmetric) design in the class of measures supported at the sets F_0 , F_{k-1} and F_k reduces to the maximization of (note that $a > b$ because otherwise by (2.12) and (4.3) we would obtain $\xi(F_{k-1}) = 0$, hence $a = b = c$, which is impossible)

$$(4.5) \quad \lambda_{\min}(M(\xi)) = \min\left\{\frac{1 + a - \sqrt{(1 - a)^2 + 4ka^2}}{2}, c - b, b\right\},$$

where $0 \leq a, b, c \leq 1$. In order to construct a good candidate, say ξ^* , for the E -optimal information matrix we assume that for the optimal design all elements in (4.5) are identical, which yields by a straightforward calculation (observing (4.3)) for the elements in the matrix (2.10)

$$(4.6) \quad a = \frac{k + 1}{k^2 + 2k + 2}, \quad b = \frac{1}{k^2 + 2k + 2}, \quad c = \frac{2}{k^2 + 2k + 2}.$$

In this case

$$(4.7) \quad \lambda_{\min}(M(\xi^*)) = \frac{1}{k^2 + 2k + 2}$$

is the minimal eigenvalue of the matrix $M(\xi^*)$ with multiplicity $s = \frac{k(k+1)}{2}$. Since this solution has been obtained under the constraint that the designs is supported at the sets F_0 , F_{k-1} and F_k and that all elements in (4.5) are identical, it is not clear that the resulting information matrix is in fact E -optimal. In a second step we establish this optimality. In order to explain the general principle we begin with an example.

Example 4.1 Consider the second order response surface model with $k = 2$ predictors. Thus, we have $m = 6$ regression functions, the minimum eigenvalue is given by $\lambda_{\min} = \frac{1}{10}$, with multiplicity $s = 3$. For a corresponding orthogonal basis in Theorem 2.1 we choose

$$q_0 = (2, -3, -3, 0, 0, 0)^T, \quad q_1 = (0, -1, 1, 0, 0, 0)^T, \quad q_2 = (0, 0, 0, 0, 0, 1)^T,$$

which yields $\|q_0\|^2 = 22$, $\|q_1\|^2 = 2$, $\|q_2\|^2 = 1$ and for the extremal polynomial

$$(4.8) \quad d(x, \varepsilon) = \frac{w_0}{\|q_0\|^2} \left(2 - 3 \sum_{j=1}^2 x_j^2 \right)^2 + \frac{w_1}{\|q_1\|^2} (x_1^2 - x_2^2)^2 + \frac{w_2}{\|q_2\|^2} (x_1 x_2)^2.$$

The vector of weights w is identified by the condition that there must be equality in (2.5) for the support points of the E -optimal design and the condition $w_0 + w_1 + w_2 = 1$. Using the points $x_{(0)} = (0, 0)^T \in F_0$ and $x_{(1)} = (1, 0)^T \in F_1$ we obtain for the vector $w = (\frac{11}{20}, \frac{3}{20}, \frac{6}{20})$ and

$$d(x, \varepsilon) = \frac{1}{10} \left(1 - 3 \left(\sum_{i=1}^2 x_i^2 \right) \left(1 - \sum_{i=1}^2 x_i^2 \right) \right) = \frac{1}{10} (1 - 3\|x\|_2^2 (1 - \|x\|_2^2)).$$

Obviously, we have for all x with $\|x\|_2 \leq 1$

$$d(x, \varepsilon) \leq \frac{1}{10} = \lambda_{\min}(M(\xi)),$$

and by Theorem 2.1 any design with information matrix of the form (2.10) with $a = \frac{3}{10}$, $b = \frac{1}{10}$, $c = \frac{2}{10}$ is E -optimal for the second order response surface model on the ball.

The following result provides a similar statement in the general case. Its proof is complicated and therefore deferred to Section 5.2.

Theorem 4.1 *Let ξ^* denote a symmetric design on the ball $\mathcal{B}_2(1)$, which puts masses*

$$(4.9) \quad \xi(F_0) = \frac{k^2}{k^2 + 2k + 2}, \quad \xi(F_{k-1}) = \frac{1}{k^2 + 2k + 2}, \quad \xi(F_k) = \frac{2k + 1}{k^2 + 2k + 2}$$

at the sets F_0 , F_{k-1} and F_k , respectively, then ξ^ is E -optimal for the second order response surface model on the k -dimensional unit ball. Moreover, the minimal eigenvalue of the matrix $M(\xi^*)$ is given by (4.7) with multiplicity $s = \frac{k(k+1)}{2}$ and the extremal polynomial in Theorem 2.1 can be chosen as*

$$(4.10) \quad d(x, \varepsilon) = \frac{1}{k^2 + 2k + 2} \left\{ 1 - \frac{2(k+1)}{k} \|x\|_2^2 (1 - \|x\|_2^2) \right\}.$$

We conclude this section with a brief discussion of rotatable designs, which are defined as designs for which the dispersion function $U : \mathcal{B}_2(1) \rightarrow \mathbb{R}; x \rightarrow U(x, \xi) = f^T(x)M^{-1}(\xi)f(x)$ is invariant with respect to orthogonal transformations, that is

$$(4.11) \quad U(x, \xi) = U(Ox, \xi), \quad \forall x \in \mathbb{R}^k,$$

whenever O is an orthogonal $k \times k$ matrix. Note that this property is equivalent to the fact that the function $U(x, \xi)$ depends only of the radius $\|x\|_2$. The following result characterizes the rotatability of a symmetric design with information matrix of the form (2.10) and will be used to investigate if E -optimal designs in the class of all rotatable designs are also E -optimal in the class of all symmetric designs.

Lemma 4.2 *Let ξ denote a symmetric design on the ball $\mathcal{B}_2(r)$ of radius $r > 0$ with information matrix of the form (2.10). Then the design ξ is rotatable for the second order response surface model, if and only if the condition*

$$(4.12) \quad c = 3b$$

is satisfied. Moreover, the uniform distribution on sphere $\partial\mathcal{B}_2(r)$ denoted by $\mathcal{U}(\partial\mathcal{B}_2(r))$ defines a rotatable design.

Proof. Let ξ denote a design with information matrix (2.10). A simple calculation shows that the inverse of the $k \times k$ upper block (3.1) of the matrix $M(\xi)$ is given by

$$M_{11}^{-1}(\xi) = \begin{pmatrix} \varkappa & q\mathbf{1}_k^T \\ q\mathbf{1}_k & G \end{pmatrix},$$

where $\varkappa = (c + b(k - 1))/Q_0$, $q = -a/Q_0$, $Q_0 = c - b + (b - a^2)k$, and $G = (d - e)I_k + e\mathbf{1}_k\mathbf{1}_k^T$ is a circulant matrix with diagonal elements d and off-diagonal elements e defined by

$$e = \frac{a^2 - b}{(c - b)Q_0}, \quad d = Q_0^{-1} - e(k - 1),$$

respectively. As a consequence, we obtain for the function U the representation

$$\begin{aligned} U(x, \xi) &= f^T(x)M^{-1}(\xi)f(x) = \varkappa + (a^{-1} + 2q)\|x\|_2^2 + (b^{-1} + 2e) \sum_{i < j}^k (x_i x_j)^2 + d \sum_{i=1}^k x_i^4 \\ &= \varkappa + (a^{-1} + 2q)\|x\|_2^2 + \left(\frac{1}{2b} + e\right)\|x\|_2^4 + \left(d - e - \frac{1}{2b}\right) \sum_{i=1}^k x_i^4. \end{aligned}$$

Now the design is rotatable if and only if the function $U(x, \xi)$ depends only on the radius $\|x\|_2$, that is

$$0 = d - e - (2b)^{-1} = (3b - c)/2b(c - b),$$

which proves the first part of the assertion. The second part follows by a straightforward calculation of the moments of the uniform distribution on the sphere $\partial\mathcal{B}_2(r)$. \square

Galil and Kiefer (1977b) have determined the E -optimal rotatable designs on the ball $\mathcal{B}_2(r)$ for the second order response surface model (2.7), which are given by

$$\xi^*(\alpha) = (1 - \alpha)\xi(\{0\}) + \alpha\mathcal{U}(\partial\mathcal{B}_2(r)),$$

where the parameter α is defined by

$$(4.13) \quad \alpha = \begin{cases} \frac{k(k+1)(k+2)}{(k+1)r^4 + k(k+2)^2}, & r^2 \leq k+2, \\ \frac{k(r^2-1)}{r^2(r^2+k-1)}, & r^2 \geq k+2. \end{cases}$$

If the design space is given by the unit ball $\mathcal{B}_2(1)$ this design is not E -optimal in the class of all designs. In fact the symmetric E -optimal design ξ^* determined in Theorem 4.1 does not satisfy condition (4.12) and is therefore not rotatable. The minimum eigenvalue of the matrix $M(\xi^*)$ is given by (4.7), while the minimum eigenvalue of the E -optimal design in the class of all rotatable designs is given by

$$\lambda_{\min}(M(\xi(\alpha))) = \frac{k+1}{k^3 + 4k^2 + 5k + 1} < \frac{1}{k^2 + 2k + 2} = \lambda_{\min}(M(\xi^*)).$$

We finally note that there exists a difference between the E - and D -optimality criterion with respect to the property of rotatability. In contrast to the E -optimal design the D -optimal design for the second order response surface model on the ball $\mathcal{B}_2(1)$ is also rotatable [see Kiefer (1960)].

Acknowledgements. This work was done during a visit of the second author at the Department of Mathematics, Ruhr-Universität Bochum, Germany. The authors would like to thank M. Stein who typed this manuscript with considerable technical expertise and F. Pukelsheim for some useful hints regarding related literature. The work of H. Dette and Y. Grigoriev was supported by the Deutsche Forschungsgemeinschaft (SFB 823: Statistik nichtlinearer dynamischer Prozesse, Teilprojekt C2).

References

- Anderson-Cook, C. M., Borror, C. M., and Montgomery, D. C. (2009). Response surface design evaluation and comparison. *Journal of Statistical Planning and Inference*, 139(2):629–641.
- Cheng, C. S. (1987). An application of the Kiefer-Wolfowitz equivalence theorem to a problem in Hadamard transform optics. *Annals of Statistics*, 15(4):1593–1603.

- Denisov, V. I. and Popov, A. A. (1976). $A - E$ -optimal and orthogonal designs of experiments for polynomial models. Preprint. Scientific Council on a Complex Problem “Cybernetic”, Academy of Science (in Russian). Moscow, Russia.
- Dette, H. (1993). A note on E -optimal designs for weighted polynomial regression. *Annals of Statistics*, 21:767–771.
- Dette, H. and Haines, L. (1994). E -optimal designs for linear and nonlinear models with two parameters. *Biometrika*, 81:739–754.
- Dette, H. and Röder, I. (1997). Optimal discrimination designs for multi-factor experiments. *Annals of Statistics*, 25:1161–1175.
- Dette, H. and Studden, W. J. (1993). Geometry of E -optimality. *Annals of Statistics*, 21:416–433.
- Draper, N. R., Heiligers, B., and Pukelsheim, F. (2000). Kiefer ordering of simplex designs for second-degree mixture models with four or more ingredients. *Annals of Statistics*, 28(2):578–590.
- Draper, N. R. and Pukelsheim, F. (2003). Canonical reduction of second-order fitted models subject to linear restrictions. *Statistics & Probability Letters*, 63(4):401–410.
- Farrel, R. H., Kiefer, J., and Walbran, A. (1967). Optimum multivariate designs. *Proc. 5th Berkeley Symp. Math. Statist. Probab.*, 1:113–138.
- Galil, Z. and Kiefer, J. (1977a). Comparison of design for quadratic regression on cubes. *J. Statist. Plann. Infer.*, 1(2):121–132.
- Galil, Z. and Kiefer, J. (1977b). Comparison of rotatable designs for quadratic regression on balls, I. *J. Statist. Plann. Infer.*, 1(1):27–40.
- Golikova, T. I. and Pantchenko, L. A. (1977). Continuous A and Q -optimal second order designs on a cube. In Nalimov, V. V., editor, *Regression Experiments (Design and Analysis) (in Russian)*, pages 71–84. Moscow University, Moscow.
- Jennrich, R. I. (1969). Asymptotic properties of non-linear least squares estimators. *Annals of Mathematical Statistics*, 40(2):633–643.
- Kiefer, J. (1959). Optimal experimental designs. *Journal of the Royal Statistical Society, Ser. B*, 21:272–319.
- Kiefer, J. (1960). Optimum experimental designs with application to systematic and rotatable designs. *Proc. 4th Berkeley Symp. Math. Statist. Probab.*, 1:381–405.
- Kiefer, J. (1961). Optimum designs in regression problems, ii. *Ann. Math. Statist.*, 32(1):298–325.
- Kiefer, J. (1974). General equivalence theory for optimum designs (approximate theory). *Annals of Statistics*, 2:849–879.
- Kiefer, J. and Wolfowitz, J. (1959). Optimum designs in regression problems. *Ann. Math. Statist.*, 30:271–294.
- Kono, K. (1962). Optimum design for quadratic regression on k -cube. *Memoirs of the Faculty of Science. Kyushu University. Series A. Mathematics*, 16:114–122.
- Laptev, V. N. (1974). *Some problems relating to construction of regression experimental designs by*

- means of computer*. PhD thesis, Novosibirsk State Technical University, Russia.
- Lim, Y. B. and Studden, J. (1988). Efficient D_s -optimal designs for multivariate polynomial regression on the q -cube. *Annals of Statistics*, 16:1225–1240.
- Melas, V. B. (1982). A duality theorem and E -optimality (translated from russian). *Industrial Laboratory*, 48:275–296.
- Melas, V. B. (2006). *Functional Approach to Optimal Experimental Design*. Springer Verlag, Berlin.
- Myers, R. H. and Montgomery, D. C. (2002). *Response Surface Methodology: Process and Product Optimization Using Designed Experiments, 2nd ed.* John Wiley & Sons, New York.
- Pázman, A. (1986). *Foundations of Optimum Experimental Design*. D. Reidel Publishing Company, Dordrecht.
- Pesotchinsky, L. L. (1975). D -optimum and quasi D -optimum second-order designs on a cube. *Biometrika*, 62:335–340.
- Pukelsheim, F. (2006). *Optimal Design of Experiments*. SIAM, Philadelphia.
- Pukelsheim, F. and Studden, W. J. (1993). E -optimal designs for polynomial regression. *Annals of Statistics*, 21(1):402–415.
- Rafajlowicz, E. and Myszk, W. (1988). Optimum experimental design for a regression on a hypercube-generalization of Hoel’s result. *Ann. Inst. Statist. Math.*, 40(4):821–827.
- Silvey, S. D. (1980). *Optimal Design*. Chapman & Hall, London.

5 Appendix: Proofs of Theorems 3.1 and 4.1

5.1 Proof of Theorem 3.1

Throughout the proof we assume $k \geq 2$, the case $k = 1$ has been treated in Pukelsheim and Studden (1993), for example. Recall the definition of the vector of regression functions (2.8) in model (2.1) and note that for the optimal design ξ^* under consideration we have $a = \frac{2}{5}$ and $b = \frac{1}{5}$ in the matrix (2.10) with minimum eigenvalue given by $\lambda_{\min}(M(\xi^*)) = \frac{1}{5}$ (see the discussion at the beginning of Section 3). Consequently, a possible candidate q_0, \dots, q_{s-1} for the basis of the eigenspace corresponding to $\lambda_{\min}(M(\xi^*))$ is given by

$$(5.1) \quad Q = (q_0, \dots, q_{s-1}) = \begin{pmatrix} G_{k \times (k+1)} & 0_{k \times k} & 0_{k \times \frac{k(k-1)}{2}} \\ 0_{\frac{k(k-1)}{2} \times (k+1)} & 0_{\frac{k(k-1)}{2} \times k} & I_{\frac{k(k-1)}{2}} \end{pmatrix}^T,$$

with an appropriate matrix $G_{k \times (k+1)} \in \mathbb{R}^{k \times k+1}$ (here and throughout this section $0_{r \times s}$ denotes the matrix with all entries given by 0). This means that the unit vectors $e_i = (0, \dots, 0, 1, 0, \dots, 0)^T$ are eigenvectors of the matrix $M(\xi^*)$ for $i = 2k + 2, \dots, m = \frac{(k+1)(k+2)}{2}$. It turns out that it is

reasonable to use a vector of weights, which is of the form

$$(5.2) \quad w = (w_0, w_1, \dots, w_{k-1}, 0, \dots, 0)^T \in \mathbb{R}^s.$$

in Theorem 2.1. Observing (5.1) it then follows that for vectors of this type only the $k + 1$ functions

$$\{1, x_1^2, \dots, x_k^2\}$$

will appear in the corresponding extremal polynomial. We now construct the remaining part of the orthogonal basis in (5.1) by choosing the block matrix

$$(5.3) \quad G_{k \times (k+1)} = \begin{pmatrix} k & -2 & -2 \mathbf{1}_{k-1}^T \\ 0_{k-1} & -\mathbf{1}_{k-1} & L \end{pmatrix} \in \mathbb{R}^{k \times (k+1)},$$

where the matrix $L = (L_{ij})_{i,j=1}^{k-1} \in \mathbb{R}^{(k-1) \times (k-1)}$ is defined by

$$L_{ij} = \begin{cases} -1 & i + j < k, \\ k - i & i + j = k, \\ 0 & i + j > k. \end{cases}$$

This gives for the eigenvectors q_0, \dots, q_{k-1} (defined by the first k rows of the matrix Q in (5.1))

$$\|q_0\|^2 = k^2 + 4k, \quad \|q_r\|^2 = (k - r)(k - r + 1), \quad r = 1, \dots, k - 1.$$

With the notation $b_i(x) = (q_i^T f(x))^2$ the extremal polynomial in Theorem 2.1 has the representation

$$(5.4) \quad d(x, \varepsilon) = \sum_{i=0}^{k-1} w_i b_i(x),$$

where we have used (5.2) and the function b_0, \dots, b_{k-1} are given by

$$(5.5) \quad \begin{aligned} b_0(x) &= \left(k - 2 \sum_{i=1}^k x_i^2 \right)^2 \cdot \frac{1}{\|q_0\|^2}, \\ b_r(x) &= \left(\sum_{i=1}^{k-r} x_i^2 - (k - r)x_{k-r+1}^2 \right)^2 \cdot \frac{1}{\|q_r\|^2}, \quad r = 1, \dots, k - 1. \end{aligned}$$

The coefficients w_i in the polynomial (5.4) are now determined by the condition $d(x, \xi) = \lambda_{\min}(M(\xi)) = \frac{1}{5}$ at the points $x^{(r)} = (0, \dots, 0, 1, \dots, 1)^T$ with $\|x^{(r)}\|_1 = r$ and the fact that $\sum_{i=0}^{k-1} w_i = 1$. This leads to the matrix equation

$$(5.6) \quad B(w_0, \dots, w_{k-1})^T = J_0,$$

where $J_0 = (\frac{1}{5}, \dots, \frac{1}{5}, 1)^T \in \mathbb{R}^k$ and the matrix $B = (B_{ir})_{i,r=0}^{k-1, k-1}$ is a lower triangular matrix with non-vanishing elements

$$B_{ir} = \begin{cases} \frac{k}{k+4} & i=0, \quad r=0, \\ \frac{(k-2i)^2}{k^2+4k} & i=1, \dots, k-2, \quad r=0, \\ \frac{(k-i)^2}{(k-r)(k-r+1)} & i=1, \dots, k-2, \quad r=1, \dots, i, \\ 1 & i=k-1, \quad r=0, \dots, k-1. \end{cases}$$

A simple calculation shows $(w_0, \dots, w_{k-1})^T = B^{-1}J_0 = \frac{1}{5k}(k+4, 4, 4, \dots, 4) \in \mathbb{R}^k$ and $w = (\frac{k+4}{5k}, \frac{4}{5k}, \dots, \frac{4}{5k}, 0, \dots, 0)^T$ is the vector which will be used for the calculation of a candidate for the extremal polynomial. For this purpose we introduce the notation

$$(5.7) \quad \alpha_r = \frac{w_r}{\|q_r\|^2} = \begin{cases} \frac{1}{5k^2} & r=0, \\ \frac{4}{5k(k-r+1)(k-r)} & r=1, \dots, k-1, \end{cases}$$

and a tedious but straightforward algebra yields for the polynomial (5.4) the representation

$$\begin{aligned} d(x, \varepsilon) &= \alpha_0 \left(k - 2 \sum_{i=1}^k x_i^2 \right)^2 + \sum_{r=1}^{k-1} \alpha_r \left(\sum_{i=1}^{k-r} x_i^2 - (k-r)x_{k-r+1}^2 \right)^2 \\ &= \frac{1}{5} \left(1 - \frac{4}{k} \sum_{i=1}^k x_i^2 (1 - x_i^2) \right), \end{aligned}$$

which coincides with (3.5). As a consequence, we obtain for all $x \in [-1, 1]^k$

$$d(x, \xi^*) \leq \lambda_{\min}(M(\xi^*)) = \frac{1}{5},$$

and by Theorem 2.1 the matrix $M(\xi^*)$ is an E -optimal information matrix. \square

5.2 Proof of Theorem 4.1.

The proof proceeds in a similar way as the proof of Theorem 3.1 but differs in some essential details from it. To be precise, recall that for the design ξ^* under consideration the minimal eigenvalue of its information matrix $M(\xi^*)$ is given by $\lambda_{\min}(M(\xi^*)) = \frac{1}{k^2+2k+2}$ and has multiplicity $s = \frac{k(k+1)}{2}$. As in the proof of Theorem 3.1 we consider the matrix defined by

(5.1) as a candidate for an orthonormal basis of the corresponding eigenspace. For the matrix $G_{k \times (k+1)} \in \mathbb{R}^{k \times k+1}$ we now use

$$(5.8) \quad G_{k \times (k+1)} = \begin{pmatrix} k & -(k+1) & -(k+1)1_{k-1}^T \\ 0_{k-1} & 1_{k-1} & L \end{pmatrix},$$

where $L = (L_{ij})_{i,j=1}^{k-1} \in \mathbb{R}^{(k-1) \times (k-1)}$ is a lower triangular matrix with non-vanishing elements

$$(5.9) \quad L_{ij} = \begin{cases} -i & i = j, \\ 1 & i > j. \end{cases}$$

Consequently, we have

$$\begin{aligned} \|q_0\|^2 &= k^2 + k(k+1)^2, \\ \|q_r\|^2 &= r(r+1), \quad r = 1, \dots, k-1, \quad \|q_r\|^2 = 1, \quad r = k, \dots, s-1, \end{aligned}$$

and with the notation $b_i(x) := (f^T(x)q_i)^2$ the candidate for the extremal polynomial in (2.5) has the representation

$$(5.10) \quad d(x, \xi^*) = \sum_{i=0}^{s-1} w_i b_i(x),$$

where (recall the definition of the vector f in (2.8))

$$\begin{aligned} b_0(x) &= \left(k - (k+1) \sum_{i=1}^k x_i^2 \right)^2 \cdot \frac{1}{\|q_0\|^2}, \\ b_r(x) &= \left(\sum_{i=1}^r x_i^2 - r x_{r+1}^2 \right)^2 \cdot \frac{1}{\|q_r\|^2}, \quad r = 1, \dots, k-1, \\ b_{k-2+i+j}(x) &= (x_i x_j)^2, \quad i = 1, \dots, k-1; \quad j = i+1, \dots, k \end{aligned}$$

(note that the eigenvectors corresponding to $b_{k-2+i+j}$ satisfy $\|q_r\| = 1$).

For determination of coordinates of the vector w we use again the fact that there must be equality in condition (2.5) of Theorem 2.1 for the support points of an E -optimal design. For the point $x_{(0)} = 0 \in \mathbb{R}^k$ the condition $d(x_{(0)}, \varepsilon) = \lambda_{\min}(M(\xi^*))$ and (5.10) then yields

$$(5.11) \quad w_0 = \frac{k^2 + 3k + 1}{k(k^2 + 2k + 2)}.$$

We now try to find a candidate for the remaining weights under the additional assumption that $p_1 := w_1 = \dots = w_{k-1}$ and $p_2 := w_k = \dots = w_{s-1}$. Because the sum of all weights is 1, this gives the equality

$$(5.12) \quad w_0 + (k-1)p_1 + \frac{k(k-1)}{2}p_2 = 1.$$

Finally, we use one more point $x_{(1)} = (1, 0, \dots, 0)^T \in F_{k-1}$ in the condition $d(x_{(1)}, \xi^*) = \lambda_{\min}(M(\xi^*))$ to obtain the equation

$$(5.13) \quad w_0 + p_1 \sum_{r=1}^{k-1} \|q_r\|^{-2} = \lambda_{\min}(M(\xi^*)) = \frac{1}{k^2 + 2k + 2}.$$

Since $\sum_{r=1}^{k-1} \|q_r\|^{-2} = 1 - k^{-1}$, we finally obtain from (5.11)-(5.13) for the weights

$$(5.14) \quad w_0 = \frac{k^2 + 3k + 1}{k(k^2 + 2k + 2)}, \quad w_1 = \dots = w_{k-1} = \frac{k + 1}{k(k^2 + 2k + 2)}, \quad w_k = \dots = w_{s-1} = \frac{2(k + 1)}{k(k^2 + 2k + 2)}.$$

Substituting these expressions in (5.10) yields by a straightforward calculation

$$(5.15) \quad d(x, \xi^*) = \frac{1}{k^2 + 2k + 2} \left(1 - \frac{2(k + 1)}{k} \|x\|_2^2 (1 - \|x\|_2^2) \right)$$

as a candidate for the extremal polynomial. Obviously, we have

$$d(x, \xi^*) \leq \frac{1}{k^2 + 2k + 2} = \lambda_{\min}(M(\xi^*))$$

for all $x \in \mathcal{B}_2(1)$, and by Theorem 2.1 the information matrix $M(\xi^*)$ defined in (2.10) with moments (4.6) is E -optimal for the second order response surface model on the ball. \square