Optimal designs for estimating the slope of a regression

Holger Dette
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum, Germany
e-mail: holger.dette@rub.de

Viatcheslav B. Melas
St. Petersburg State University
Department of Mathematics
St. Petersburg, Russia
email: v.melas@pochta.tvoe.tv

Andrey Pepelyshev
St. Petersburg State University
Department of Mathematics
St. Petersburg , Russia
email: andrey@ap7236.spb.edu

October 7, 2008

Abstract

In the common linear regression model we consider the problem of designing experiments for estimating the slope of the expected response in a regression. We discuss locally optimal designs, where the experimenter is only interested in the slope at a particular point, and standardized minimax optimal designs, which could be used if precise estimation of the slope over a given region is required. General results on the number of support points of locally optimal designs are derived if the regression functions form a Chebyshev system. For polynomial regression and Fourier regression models of arbitrary degree the optimal designs for estimating the slope of the regression are determined explicitly for many cases of practical interest.

Keyword phrases: locally optimal design, standardized minimax optimal design, estimating derivatives, polynomial regression, Fourier regression

AMS Subject Classification: Primary 62K05

1 Introduction

Consider the common linear regression model

$$Y_i = \theta^T f(x_i) + \varepsilon_i, \qquad i = 1, \ldots, N \tag{1.1}$$

where $\theta \in \mathbb{R}^m$ denotes the vector of unknown parameters, $f(x) = (f_1(x), \ldots, f_m(x))^T$ is the vector of regression functions and x varies in the design space $\mathcal{X} = [a, b]$ $(a < b)$. In (1.1) $\varepsilon_1, \ldots, \varepsilon_n$ denote uncorrelated random variables with $\mathrm{E}[\varepsilon_i] = 0; \mathrm{Var}(\varepsilon_i) = \sigma^2 > 0$ $(i = 1, \ldots, N)$ and it is assumed that the regression functions are differentiable on $\mathcal{Z} = [a, b] \cup [a', b']$, where $[a', b']$ is an interval of interest that can coincide with $[a, b]$ or intersect it.

Most work on optimal experimental design for the regression model (1.1) refers to precise estimation of the vector of unknown parameters or to the estimation of the expected response $\theta^T f(x)$ [see e.g. Silvey (1980) or Pukelsheim (1993)]. However in many experiments differences in the response will often be of more importance than the absolute response. If one is interested in a difference at two points close together, this means that the estimation of the local slope of the expected response is the main object of inference of the experiment.

The present paper is devoted to the problem of optimal designing experiments for estimating the slope of the expected response in a regression. Pioneering work in this direction has been done by Atkinson (1970) and the problem has subsequently been taken up by many other authors [see e.g. Ott and Mendenhall (1972), Murthy and Studden (1972), Myres and Lahoda (1975), Hader and Park (1978), Mukerjee and Huda (1985), Mandal and Heiligers (1992), Pronzato and Walter (1993) and Melas et al. (2003)]. The present paper takes a closer look at design problems of this type in the context of a one-dimensional predictor, in particular for polynomial regression and trigonometric regression models. In Section 2 we introduce two optimal design problems which might be appropriate if one goal of the experiment consists in the estimation of the slope of a regression, a locally and a standardized minimax optimality criterion. While the locally optimal design refers to the estimation of the slope at a particular point, the standardized minimax optimal design is appropriate, if the experimenter is interested in the estimation of the slope over a certain range. We state a general result regarding the number of support points of locally optimal designs for estimating the slope if the regression functions form a Chebyshev system. Section 3 discusses the polynomial regression model in more detail. In particular, it is shown that the optimal design problem for estimating the slope of the expected response in a polynomial regression can be reduced to the problem of estimating individual coefficients in this model [see Sahm (1998) or Dette et al. (2004)]. Using these results it is possible to derive a rather complete and explicit description of the optimal designs for the estimation of the slope for arbitrary degree of the polynomial. The case of trigonometric regression is discussed in Section 4. We explicitly determine locally and standardized maximin optimal designs for estimating the slope of the expected response in these models. Finally, all technical proofs are deferred to an Appendix in Section 5.

2 Optimal designs for estimating the slope

Consider the linear regression model (1.1). An approximate design is a probability measure on the interval $[a, b]$ with finite support [see e.g. Kiefer (1974)], say

$$\xi = \begin{pmatrix} x_1 & \dots & x_k \\ w_1 & \dots & w_k \end{pmatrix}, \tag{2.1}$$

where the support points $x_1, \dots, x_k$ give the positions in the interval $[a, b]$ at which observations are taken and the weights give the relative proportions of total observations taken at the corresponding support points. If N observations can be performed by the experimenter, a rounding procedure is applied to obtain the samples sizes $N_i \approx w_i N$ at the experimental conditions x_i, $i = 1, 2, \dots, k$, subject to $N_1 + N_2 + \dots + N_k = N$ [see Pukelsheim and Rieder (1992)]. In this case the covariance matrix of the least squares estimate $\hat{\theta}$ is approximately given by the matrix $\frac{\sigma^2}{N} M^{-1}(\xi)$, where

$$M(\xi) = \int_a^b f(x) f^T(x) d\xi(x) \tag{2.2}$$

denotes the information matrix of the design ξ. If the estimation of the expected response $\theta^T f(x)$ or the parameter θ is the main goal of the experiment, an optimal design minimizes (or maximizes) a specific convex (or concave) function of the information matrix and there are numerous optimality criteria proposed in the literature, which can be used for the determination of efficient designs [see e.g. Silvey (1980) or Pukelsheim (1993)]. Note that the least squares estimate for the slope of the expected response $\theta^T f'(x) = \sum_{j=1}^m \theta_j f_j'(x)$ at the point $x \in \mathcal{X}$ is given by $\hat{\theta}^T f'(x)$. Consequently, if the data is collected according to an approximate design, then the variance of $\hat{\theta}^T f'(x)$ is approximately given by

$$\mathrm{Var}(\hat{\theta}^T f'(x)) \approx \frac{\sigma^2}{N} \Phi(\xi), \tag{2.3}$$

where

$$\Phi(\xi) = \begin{cases} (f'(x))^T M^-(\xi) f'(x) & \text{if} \quad f'(x) \in \text{Range } (M(\xi)) \\ \infty & \text{else} \end{cases} \tag{2.4}$$

Therefore a design, say ξ_x^*, minimizing $\Phi(\xi)$ in the class of all (approximate) designs satisfying $f'(x) \in \text{Range } (M(\xi))$ is called locally optimal design for estimating the slope of the expected response, where the term "locally" reflects the fact that we are minimizing the variance of the least squares estimate of the slope of the expected response at the particular point $x \in \mathcal{X}$. On the other hand, if the interest of the experimenter lies in the slope of the expected response over a certain region, say $[a', b']$, a maximin approach might be more appropriate [see e.g. Dette (1995), Müller (1995) or Müller and Pázman (1998)]. To be precise, recall that ξ_x^* denotes the locally optimal design for estimating the slope of the expected response at the point x, then we call a design ξ^* standardized minimax or minimax efficient optimal design for the estimation of the slope of the expected response if ξ^* minimizes the expression

$$\text{eff}(\xi) = \max_{x\in[a',b']} \frac{(f'(x))^T M^-(\xi) f'(x)}{(f'(x))^T M^-(\xi_x^*) f'(x)} \in [1,\infty). \tag{2.5}$$

In nearly all cases of practical interest, standardized optimal designs have to be determined numerically [see for example Müller (1995), Dette et al. (2003) and Dette and Braess (2007)]. Note also that the calculation of the standardized optimal designs for the estimation of the slope of the expected response requires the determination of the locally optimal designs, for which analytical results will be derived in the following sections in the case of a polynomial and the trigonometric regression model (for the last named model we also derive explicit standardized minimax optimal designs). Before we consider these special cases, we present a general result regarding the number of support points of locally optimal designs if the regression functions form an extended Chebyshev system of second order. Recall that the functions $f_1(x), \dots, f_m(x)$ generate an extended Chebyshev system of order 2 on the set $\mathcal{Z} = [a,b] \cup [a',b']$ if and only if

$$U^* \begin{pmatrix} 1 & \dots & m \\ x_1 & \dots & x_m \end{pmatrix} > 0$$

for all $x_1 \le \dots \le x_m$ ($x_j \in \mathcal{X}$; $j = 1,\dots,m$) where equality occurs at at most 2 consecutive points x_j, the determinant U^* is defined by

$$U^* \begin{pmatrix} 1 & \dots & m \\ x_1 & \dots & x_m \end{pmatrix} = \det(f(x_1),\dots,f(x_m))$$

and the columns $f(x_i), f(x_{i+1})$ are replaced by $f(x_i), f'(x_{i+1})$ if the points x_i and x_{i+1} coincide. Note that under this assumption any linear combination

$$\sum_{i=1}^m \alpha_i f_i(x)$$

($\alpha_1,\dots,\alpha_m \in \mathbb{R}, \sum_{i=1}^m \alpha_i^2 \neq 0$) has at most $m-1$ roots, where multiple roots are counted twice [see Karlin and Studden (1966), Ch. 1]. Because $\{f_1(x),\dots,f_m(x)\}$ is also a Chebyshev system on the interval $[a,b]$, it follows that there exist m points, say $a \le x_1^* < \dots < x_m^* \le b$ and coefficients $\alpha_1^*,\dots,\alpha_m^*$ such that the "polynomial" $\alpha^{*T} f(x) = \sum_{j=1}^m \alpha_j^* f_j(x)$ satisfies

$$|\alpha^{*T} f(x)| \le 1 \qquad \forall x \in [a,b] \tag{2.6}$$

$$\alpha^{*T} f(x_j^*) = (-1)^j \qquad j = 1,\dots,m. \tag{2.7}$$

The function $\alpha^{*T} f(x)$ (which is not necessarily unique) is called (generalized) Chebyshev polynomial, while the points $x_1^*,\dots,x_m^*$ are called Chebyshev points. Note that the Chebyshev polynomial and the Chebyshev points are determined uniquely under the condition that the constant function is an element of $span(f_1, \ldots, f_m)$ [see Karlin and Studden (1966), Ch. 1]. The following result specifies the number of support points of the locally optimal design for estimating the slope of the expected response, if the regression functions form a Chebyshev system. The proof is deferred to the Appendix.

Theorem 2.1. *Assume that the regression functions in model* (1.1) *form an extended Chebyshev system of second order on the interval* $[a, b]$, *then the number of support points of any locally optimal design for estimating the slope of the expected response is at least* $m - 1$. *Moreover, if the number of support points is* m, *then these points must be Chebyshev points. If the constant function is an element of* $span(f_1, \ldots, f_m)$ *then the number of support points is at most* m.

3 Optimal designs for estimating the slope of a polynomial regression

It is well known that the functions $f_1(x) = 1, f_2(x) = x, \ldots, f_m(x) = x^{m-1}$ form an extended Chebyshev system of order two on any arbitrary nonnegative interval [see Karlin and Studden (1966)]. For this choice the model (1.1) reduces to the common polynomial regression

$$Y_i = \sum_{j=1}^{m} \theta_j x_i^{j-1} + \varepsilon_i; \qquad i = 1, \ldots, N, \tag{3.1}$$

for which locally optimal designs for estimating the slope of the expected response have been discussed by Murthy and Studden (1972) for the quadratic and cubic model. In this section we will derive a general solution of this design problem for any $m \geq 3$ reducing the optimization to a design problem for estimating individual coefficients as considered by Sahm (1998) and Dette et al. (2004). For this purpose we denote by $e_j = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ the jth unit vector in $\mathbb{R}^m$ and call a design minimizing $e_j^T M^-(\xi) e_j$ in the class of all designs satisfying $e_j \in \text{Range}(M(\xi))$ an e_j-optimal design or optimal design for estimating the coefficient θ_j in the polynomial regression model (3.1). The following result relates an optimal design problem for estimating the slope of the expected response in a polynomial regression to an e_j-optimal design problem and is proved in the Appendix.

Theorem 3.1. *A design*

$$\xi = \begin{pmatrix} x_1 & \ldots & x_k \\ w_1 & \ldots & w_k \end{pmatrix}$$

is locally optimal for estimating the slope of expected response at the point x in the polynomial regression model (3.1) *on the interval* $[a,b]$ *if and only if the design*

$$\eta = \begin{pmatrix} x_1 - x & \dots & x_k - x \\ w_1 & \dots & w_k \end{pmatrix} \tag{3.2}$$

is an e_2-optimal for the polynomial regression model (3.1) *on the interval* $[a-x, b-x]$.

We can now apply the results of Sahm (1998) and Dette et al. (2004) to derive the locally optimal designs for estimating the slope of the expected response in the polynomial regression model (3.1). To be precise, we consider the sets

$$\begin{aligned} A_i &= (-\nu_{m-1-i}, \nu_{i+1})\,; & i &= 0, \dots, m-2, \\ B_{1,i} = -B_{2,i} &= [\nu_i, \rho_i]\,; & i &= 1, \dots, m-2, \\ C_i &= (\rho_i, -\rho_{m-1-i})\,; & i &= 1, \dots, m-2, \end{aligned} \tag{3.3}$$

where $\nu_{m-1} = \infty$ and $\nu_1 < \nu_2 < \dots, \nu_{m-2}$ are the roots of the first derivative of the polynomial

$$(x+1)U_{m-2}(x),$$

and $U_j(x) = \sin((j+1)\arccos x)/\sin(\arccos x)$ is the j-th Chebyshev polynomial of the second kind [see Szegö (1975)]. The points ρ_i are obtained from these roots via the transformation

$$\rho_i = \nu_i + (1+\nu_i)\frac{1-\cos(\pi/(m-1))}{1+\cos(\pi/(m-1))}.$$

Define $u_i = \cos(\pi\frac{i-1}{m-1})$ $(i = 1, \dots, m)$ as the extreme points of the mth Chebyshev polynomial of the first kind $T_n(x) = \cos((m-1)\arccos x)$, $x_i = \frac{b-d}{2}u_i + \frac{b+d}{2}$ $(i = 1, \dots, m)$ and set

$$\left.\begin{aligned} \hat{x}_{ij} &= \frac{b-a}{2}\,\hat{b}_{ij} + \frac{b+a}{2}, \\ \overline{x}_{ij} &= \frac{b-a}{2}\,\overline{b}_{ij} + \frac{b+a}{2}, \end{aligned}\right. \qquad i = 1, \dots, m;\ j = 1, \dots, m-2\ ;$$

where

$$\left.\begin{aligned} \hat{b}_{ij} &= -\frac{(1+x)(v_j - u_i)}{v_{j+1}} + x, \\ \overline{b}_{ij} &= \frac{(1-x)(v_j + u_i)}{v_{j+1}} + x, \end{aligned}\right. \qquad i = 1, \dots, m\ ; j = 1, \dots, m-2\ .$$

Finally, let $\mathrm{L}_1(x), \dots, \mathrm{L}_m(x)$, $\hat{\mathrm{L}}_{1j}(x), \dots, \hat{\mathrm{L}}_{mj}(x)$ and $\overline{\mathrm{L}}_{1j}(x), \dots, \overline{\mathrm{L}}_{mj}(x)$ denote the Lagrange interpolation polynomials with knots $x_1^*, \dots, x_m^*$; $\overline{x}_{1j}, \dots, \overline{x}_{mj}$ and $\hat{x}_{1j}, \dots, \hat{x}_{mj}$, respectively ($j =$

$1,\dots,m-2)$, then we consider the designs

$$(3.4) \qquad \xi^* = \begin{pmatrix} x_1^* & \dots & x_m^* \\ w_1^* & \dots & w_m^* \end{pmatrix}$$

$$(3.5) \qquad \hat{\xi}_j = \begin{pmatrix} \hat{x}_{1j} & \dots & \hat{x}_{m-1,j} \\ \hat{w}_{1j} & \dots & \hat{w}_{m-1,j} \end{pmatrix} \; ; \quad j=1,\dots,m-2,$$

$$(3.6) \qquad \overline{\xi}_j = \begin{pmatrix} \overline{x}_{1j} & \dots & \overline{x}_{m-1,j} \\ \overline{w}_{1j} & \dots & \overline{w}_{m-1,j} \end{pmatrix} \; ; \quad j=1,\dots,m-2,$$

where $x_i^* = \frac{b-a}{2}\cos(\frac{i-1}{m-1}\pi) + \frac{a+b}{2}$ $(i=1,\dots,m-1)$ are the Chebyshev points on the interval $[a,b]$ and the weights are given by

$$(3.7) \qquad w_i^* = \frac{|\mathrm{L}_i'(x)|}{\sum_{k=1}^m |\mathrm{L}_k'(x)|} \; ; \quad i=1,\dots,m,$$

$$(3.8) \qquad \hat{w}_{ij} = \frac{|\hat{\mathrm{L}}_{ij}'(x)|}{\sum_{k=1}^m |\hat{\mathrm{L}}_{kj}'(x)|} \; ; \quad i=1,\dots,m; \; j=1,\dots,m-2,$$

$$(3.9) \qquad \overline{w}_{ij} = \frac{|\overline{\mathrm{L}}_{ij}'(x)|}{\sum_{k=1}^m |\overline{\mathrm{L}}_{kj}'(x)|} \; ; \quad i=1,\dots,m; \; j=1,\dots,m-2,$$

respectively (note that $\overline{w}_{mj} = \hat{w}_{mj} = 0$). An application of the results of Sahm (1998) and Dette et al. (2004) now yields the following result.

Corollary 3.2. *For each $x \in (-\infty,\infty)$ the locally optimal design for estimating the slope of the expected response at the point x in the polynomial regression model* (3.1) *is unique.*

(i) If

$$\frac{a+b-2x}{b-a} \in \bigcup_{i=1}^{m-2} A_i,$$

then the optimal design is given by the design ξ^* defined in (3.4).

(ii) If for some $(j=1,\dots,m-2)$

$$\frac{a+b-2x}{b-a} \in B_{1,j},$$

then the optimal design is given by the design $\hat{\xi}_j$ defined in (3.5).

(iii) If for some $j=1,\dots,m-2$

$$\frac{a+b-2x}{b-a} \in B_{2,j},$$

then the optimal design is given by the design $\overline{\xi}_j$ defined in (3.6).

(iv) If for some $j = 1, \dots, m-2$

$$\frac{a+b-2x}{b-a} \in \bigcup_{i=0}^{m-2} C_j, \tag{3.10}$$

then the optimal design is supported at $m-1$ points including the boundary points of the interval $[a, b]$.

Remark 3.3. If condition (3.10) is satisfied, the optimal design for estimating the expected response of the polynomial regression cannot be found explicitly. Dette et al. (2004) provided a numerical procedure based on the implicit function theorem for the construction of e_k-optimal designs in polynomial regression models, which can easily be adapted to the problem of designing experiments for estimating the slope of the expected response in a polynomial regression. The details are omitted for the sake of brevity.

Example 3.4. Consider the case of a quadratic regression, that is $m = 3$, and $[a, b] = [-1, 1]$. In this case we have

$$\begin{aligned} \nu_1 &= -\frac{1}{2}, \ \nu_2 = \infty, \qquad \rho_1 = 0\,; \\ A_0 &= \Big(-\infty, -\frac{1}{2}\Big), \quad A_1 = \Big(\frac{1}{2}, \infty\Big)\,; \\ B_{1,1} &= \Big(-\frac{1}{2}, 0\Big], \quad B_{2,1} = \Big[0, \frac{1}{2}\Big), \quad C_1 = \emptyset. \end{aligned}$$

By Corollary 3.2 the unique optimal design for estimating the slope of the quadratic regression is given by

$$\xi^* = \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{4} - \frac{1}{8}x & \frac{1}{2} & \frac{1}{4} + \frac{1}{8}x \end{pmatrix}$$

if $x \in A_1 \cup A_2 = \mathbb{R} \setminus [-\frac{1}{2}, \frac{1}{2}]$, by

$$\hat{\xi} = \begin{pmatrix} 2x-1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

if $x \in B_{1,1} = [-\frac{1}{2}, 0]$ and by

$$\bar{\xi} = \begin{pmatrix} -1 & 1-2x \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

if $x \in B_{2,1} = [0, \frac{1}{2}]$. This case was obtained independently by a direct calculation in Fedorov and Müller (1997).

Example 3.5. For the cubic regression model, that is $m = 4$, on the interval $[-1, 1]$ we have

$$\begin{array}{lll} \nu_1 = -\frac{1}{3} - \frac{\sqrt{7}}{6}, & \nu_2 = -\frac{1}{3} + \frac{\sqrt{7}}{6}, & \nu_3 = \infty \\ \rho_1 = \frac{-1-2\sqrt{7}}{9}, & \rho_2 = \frac{-1+2\sqrt{7}}{9} & \end{array}$$

which gives

$$\begin{aligned} A_0 &= \Big(-\infty, -\frac{1}{3} - \frac{\sqrt{7}}{6}\Big),\ A_1 = \Big(\frac{1}{3} - \frac{\sqrt{7}}{6}, -\frac{1}{3} + \frac{\sqrt{7}}{6}\Big),\ A_2 = \Big(\frac{1}{3} + \frac{\sqrt{7}}{6}, \infty\Big)\ ; \\ B_{1,1} &= \Big[-\frac{1}{3} - \frac{\sqrt{7}}{6}, -\frac{1}{9} - \frac{2\sqrt{7}}{9}\Big],\ B_{1,2} = \Big[-\frac{1}{3} + \frac{\sqrt{7}}{6}, -\frac{1}{9} + \frac{2\sqrt{7}}{9}\Big]\ ; \\ B_{2,1} &= \Big[\frac{1}{9} + \frac{2\sqrt{7}}{9}, \frac{1}{3} + \frac{\sqrt{7}}{6}\Big],\ B_{2,2} = \Big[\frac{1}{9} - \frac{2\sqrt{7}}{9}, \frac{1}{3} - \frac{\sqrt{7}}{6}\Big]\ ; \\ C_1 &= \Big(-\frac{1}{9} - \frac{2\sqrt{7}}{9}, \frac{1}{9} - \frac{2\sqrt{7}}{9}\Big),\ C_2 = \Big(-\frac{1}{9} + \frac{2\sqrt{7}}{9}, \frac{1}{9} + \frac{2\sqrt{7}}{9}\Big). \end{aligned}$$

Consequently, if $x \in A_1 \cup A_2 \cup A_3$ the optimal design for estimating the slope of the expected response in the polynomial regression is supported at the Chebyshev points $x_1^* = -1, x_2^* = -\frac{1}{2}, x_3^* = \frac{1}{2}, x_4^* = 1$ with weights given in (3.7). If $x \in B_{1,1}$ then the locally optimal design for estimating the slope at the point x has three support points, i.e.

$$x_1^* = -1, x_2^* = \frac{-1 + \sqrt{7} + 3x}{4 - \sqrt{7}}, x_3^* = \frac{5 + \sqrt{7} + 9x}{4 - \sqrt{7}}.$$

If $x \in B_{1,2}$ then the support points of the locally optimal design are given by

$$x_1^* = -1, x_2^* = \frac{-1 - \sqrt{7} + 3x}{4 + \sqrt{7}}, x_3^* = \frac{5 - \sqrt{7} + 9x}{4 + \sqrt{7}}.$$

If $x \in B_{2,1}$ then the locally optimal design has 3 support points, i.e.

$$x_1^* = \frac{-5 - \sqrt{7} + 9x}{4 - \sqrt{7}}, x_2^* = \frac{1 - \sqrt{7} + 3x}{4 - \sqrt{7}}, x_3^* = 1.$$

Finally, if $x \in B_{2,2}$ then

$$x_1^* = \frac{-5 + \sqrt{7} + 9x}{4 + \sqrt{7}}, x_2^* = \frac{1 + \sqrt{7} + 3x}{4 + \sqrt{7}}, x_3^* = 1.$$

These results have also been obtained independently by Murthy and Studden (1972) by a direct calculation. If $x \in C_1 \cup C_2$ then there is no explicit form, but the optimal design for estimating the slope of the expected response at the point x is supported at the two boundary points and a third point, which can be determined by means of a Taylor expansion as described in Dette et al. (2004).

We conclude this section with a brief discussion of the standardized minimax designs in the quadratic and cubic regression model which have to be found numerically in nearly all cases of practical interest. Here we take $[a', b'] = [a, b]$. Müller and Pázman (1998) determined the standardized minimax optimal design for estimating the slope in the quadratic regression explicitly as

$$\xi^*_{st} = \begin{pmatrix} -1 & 0 & 1 \\ \alpha/2 & 1-\alpha & \alpha/2 \end{pmatrix} \tag{3.11}$$

where $\alpha = 23 - 10\sqrt{5} \approx 0.64$. A numerical calculation shows that the standardized minimax optimal design for estimating the slope in the cubic regression is given

$$\xi^*_{st} = \begin{pmatrix} -1 & -0.39 & 0.39 & 1 \\ 0.23 & 0.27 & 0.27 & 0.23 \end{pmatrix}. \tag{3.12}$$

We note that the standardized minimax optimal designs (3.11) and (3.12) are very similar to the D-optimal designs for a quadratic or cubic regression on the interval $[-1, 1]$, respectively.

4 Optimal design for estimating slopes in Fourier regression

In the context of trigonometric regression models

$$Y_i = \theta_0 + \theta_1 \sin x_i + \theta_2 \cos x_i + \cdots + \theta_{2k-1} \sin(kx_i) + \theta_{2k} \cos(kx_i) + \varepsilon_i \ , \quad i = 1, \ldots, N \tag{4.1}$$

the locally and standardized minimax optimal design for estimating the slope of the expected response can be found analytically in most cases, if the design space is given by the interval $[0, 2\pi)$. These models are widely used to describe periodic phenomena [see e.g. Lestrel (1997), Lau and Studden (1985), Wu (2002) or Zen and Tsai (2004) among others] and optimal design problems for estimating the parameter $\theta = (\theta_0, \ldots, \theta_{2k})$ have been discussed by several authors. For the determination of optimal design for estimating the slope of the expected response in the Fourier regression model (4.1) we note that the functions $1, \sin x, \cos x, \ldots, \sin(kx), \cos(kx)$ form a Chebyshev system on the interval $[0, 2\pi)$ and define $[x + t]_+ = x + t + 2\pi s$, where s is the (unique) integer such that $x + t + 2\pi s \in [0, 2\pi)$.

Theorem 4.1. *The locally optimal design for estimating the slope of the expected response in the trigonometric regression model* (4.1) *is given by*

$$\xi^*_x = \begin{pmatrix} [x_1 - x]_+ & \ldots & [x_{2k} - x]_+ \\ w_1 & \ldots & w_{2k} \end{pmatrix}, \tag{4.2}$$

where $x_i = \pi(2i-1)/2k$ $(i = 1, \ldots, 2k)$, *and weights are defined by*

$$w_i = \frac{|A_i|}{\sum_{j=1}^{2k} |A_j|} \qquad i = 1, \ldots, 2k$$

$A_i = e_i^T F^{-1} f'_-(0)$, $F = (f_i(x_j)(-1)^j)_{i,j=1}^{2k}$ *and* $f_-(x) = (1, \sin x, \cos x, \ldots, \sin((k-1)x), \cos((k-1)x), \sin(kx))^T$. *Moreover,* $\Phi(\xi_x^*) = k^2$.

Theorem 4.2. *Any equispaced design* ξ^* *with* $n \geq 2k+1$ *support points is standardized minimax optimal for estimating the slope of the expected response in the trigonometric regression model* (4.1). *Moreover, the maximal efficiency is given by*

$$c_k(\xi^*) = \frac{3k}{(k+1)(2k+1)}.$$

5 Appendix: Proofs.

Let $c \in \mathbb{R}^m$ and recall that a c-optimal design minimizes the expression $c^T M^-(\xi)c$ in the class of all designs for which $c \in \text{Range}(M(\xi))$. Note that the choice $c = f'(x)$ yields the locally optimal design problem for the estimation of the slope of the expected response. The following result is a reformulation of the equivalence theorem for c-optimality [see Pukelsheim (1993)].

Lemma 5.1. *If* $f_1, \ldots, f_m$ *are continuous functions and form a Chebyshev system on the interval* $[a, b]$, *then the design* ξ *given by* (2.1) *is c-optimal if and only if there exists a vector* $q \in \mathbb{R}^m$, *such that the generalized polynomial* $q^T f(x)$ *satisfies the following conditions*

(i) $q^T f(x_i) = (-1)^i \qquad i = 1, \ldots, m$

(ii) $|q^T f(x)| \leq 1 \qquad$ *for all* $\quad x \in [a, b]$

(iii) $Fw = hc$

for some $h > 0$, *where* $F = (f_i(x_j))_{i,j=1}^{m,k}$ *and* $w = (w_1, \ldots, w_k)$. *Moreover,* $c^T M^-(\xi)c = 1/h^2$.

Proof of Theorem 2.1. Assume that ξ is a locally optimal design for estimating the slope of the expected response at the point x which has the form (2.1), where the number of support points satisfies $k \leq m-2$. We only consider the case $x \notin \text{supp}(\xi)$ and $k = m-2$ (the case $k < m-2$ or $x \in \text{supp}(\xi)$ is treated similarly). From Lemma 5.1 it follows for the vector $\alpha = (w_1, \ldots, w_{m-2}, 0) \in \mathbb{R}^{m+1}$ and the matrix $\tilde{F} = (f(x_1)(-1), f(x_2), \ldots, f(x_{m-2})(-1)^{m-2}, f(x)) \in \mathbb{R}^{m \times m-1}$ that

$$\tilde{F}\alpha = h \; f'(x),$$

which implies

$$\det\,(f(x_1),\ldots,f(x_{m-2}),f(x),f'(x))=0.$$

However this condition contradicts the property that the functions $f_1,\ldots,f_m$ generate an extended Chebyshev system of second order on the set $[a,b]$. Consequently (using similar arguments for the other cases), it follows that $k\geq m-1$. If $k=m$ we obtain from part (i) and (ii) of Lemma 5.1 that $x_1,\ldots,x_m$ are Chebyshev points. Now, let us assume that the constant function is an element of $span(f_1,\ldots,f_m)$. Then the points with properties (i) and (ii) are uniquely determined. It means that there are no optimal designs with $k>m$ support points.

□

Proof of Theorem 3.1. Obviously we have

$$f(x+t)=\begin{pmatrix}1\\ x+t\\ \vdots\\ (x+t)^{m-1}\end{pmatrix}=L_xf(t), \tag{A.1}$$

where L_x is a lower triangular matrix which does not depend on t and $f(t)=(1,t,\ldots,t^{m-1})^T$. Consequently it follows that

$$\frac{\partial}{\partial t}f(x+t)=L_xf'(t),$$

which implies $f'(x)=L_xf'(0)$. This yields for the information matrices of the designs ξ and η

$$M(\xi)=\sum_{i=1}^k w_if(x_i)f^T(x_i)=L_x\left\{\sum_{i=1}^k w_if(x_i-x)f^T(x_i-x)\right\}L_x^T=L_xM(\eta)L_x^T, \tag{A.2}$$

and for the criterion $\Phi(\xi)$

$$\Phi(\xi)=(f'(x))^TM^-(\xi)f'(x)=(f'(0))^TM^-(\eta)f'(0). \tag{A.3}$$

If ξ is locally optimal for estimating the slope of the expected response in the polynomial regression model (3.1) it follows from the identity (A.3) that the corresponding design η defined by (3.2) is e_2-optimal for the polynomial regression model (3.1) on the interval $[a-x,b-x]$ and vice versa, which proves the assertion of Theorem 3.1. □

Proof of Theorem 4.1. Let ξ denote a design with masses $w_1,\ldots,w_n$ at the points $x_1,\ldots,x_n$ and η a design with the same masses at the points $x_1-x,\ldots,x_n-x$, then a straightforward calculation yields

$$\int_0^{2\pi}f(x)f^T(x)d\xi(x)=B\int_0^{2\pi}f(x)f^T(x)d\eta(x)B^T,$$

where B is a $(2k+1) \times (2k+1)$ block diagonal matrix with blocks $1, B_1, \ldots, B_k$, such that

$$B_l = \begin{pmatrix} \cos(lx) & -\sin(lx) \\ \sin(lx) & \cos(lx) \end{pmatrix}. \tag{A.4}$$

Consequently it follows from the identity $(f'(x))^T B^{-1} = (0, 1, 0, 2, 0, \ldots, k, 0)^T \in \mathbb{R}^{2k+1}$ that the locally optimal design for estimating the slope of the expected response in the trigonometric regression at the point x can be obtained from the locally optimal design for estimating the slope at the point 0 using the transformation $x_i \to x_i - x \pmod{2\pi}$. A straightforward calculation shows that for any α the function

$$\cos[k(x-\alpha)] = \cos(k\alpha)\cos(kx) + \sin(k\alpha)\sin kx$$

is the unique trigonometric polynomial of degree k, which attains the value 1 at the point α and attains its maximal absolute value 1 over the interval $[0, 2\pi)$ at $n \geq 2k$ points. It therefore follows from Lemma 5.1 that the locally optimal design ξ_0^* for estimating the slope of the trigonometric regression at the point 0 is supported at the points $x_i = [i\pi/k + \alpha]_+$ $(i = 1, \ldots, 2k)$ for some $\alpha \in \mathbb{R}$. We define $\overline{\xi}_0$ as the design with the same masses as the design ξ_0^* at the points $\overline{x}_i = [2\pi - x_i]_+$ $(i = 1, \ldots, 2k)$, then we obtain from Lemma 5.1 that the design $\overline{\xi}_0$ is also locally optimal for estimating the slope of the expected response in the trigonometric regression at the point 0. Consequently a further optimal design is given by the convex combination $\frac{1}{2}(\xi_0^* + \overline{\xi}_0)$. Because this design has at most $2k+1$ support points (by Lemma 5.1) it follows that $\xi_0^* = \overline{\xi}_0$, $\alpha = \frac{\pi}{2k}$, which implies $x_i = (2i-1)\pi/(2k)$ $(i = 1, \ldots, 2k)$. The formula for the weights is obtained from Lemma 5.1, while a direct calculation shows that the quantity h in Lemma 5.1 is given by $h = 1/k$, which implies $\Phi(\xi) = k^2$. □

Proof of Theorem 4.2. Let ξ^* denote an equispaced design with $n \geq 2k+1$ support points. It follows from Pukelsheim (1993) that the information matrix of the design ξ^* is given by $M(\xi^*) = \text{diag}(1, 1/2, \ldots, 1/2) \in \mathbb{R}^{2k+1 \times 2k+1}$ and the efficiency for estimating the slope at the point x is given by $c_k = 3k/((k+1)(2k+1))$ for any x. Assume that $b' - a' \geq 2\pi(1 - 1/(2k+1))$ and assume that there exists a design $\overline{\xi}$ with

$$\max_{x \in [a', b']} \frac{(f'(x))^T M^-(\overline{\xi}) f'(x)}{k^2} < \frac{1}{c_k}.$$

Now consider the matrix $A = \text{diag}(0, 1, 1, \ldots, k^2, k^2)$ and the design

$$\eta^* = \begin{pmatrix} x_1^* + a' & \ldots & x_{2k+1}^* + a' \\ \frac{1}{2k+1} & \ldots & \frac{1}{2k+1} \end{pmatrix}$$

with $x_i^* = (i-1)2\pi/(2k+1)$ $i = 1, \ldots, 2k+1$. We obtain by a direct calculation that η^* is a D-optimal design for the trigonometric regression model on the interval $[a', a' + 2\pi]$, that is $M(\eta^*) = \text{diag}(1, 1/2, \ldots, 1/2)$, and therefore

$$\int f'(x)(f'(x))^T d\eta^*(x) = (1/2)A.$$

This yields

$$\text{(A.5)}\quad \frac{1}{c_k} > \frac{1}{k^2}\max_{x\in[a',b']}(f'(x))^T M^{-1}(\bar{\xi})f'(x) \geq \frac{1}{k^2}\mathrm{tr}\left\{M^{-1}(\bar{\xi})\int f'(x)(f'(x))^T d\eta^*(x)\right\}$$
$$= \frac{1}{2k^2}\mathrm{tr}(M^{-1}(\bar{\xi})A) \geq \frac{1}{k^2}\mathrm{tr}(A) = \frac{1}{c_k}.$$

Here the last inequality follows from the fact that the design ξ^* minimizes the function $\mathrm{tr}(M^{-1}(\bar{\xi})A)$ in the class of all approximate designs with minimal value $2\mathrm{tr}(A)$ [this can be proved by a standard application of the equivalence theorem for linear optimality criteria; see Fedorov (1972)]. Because of the contradiction in (A.5) the assertion of Theorem 4.2 follows. □

Acknowledgements. The authors are grateful to Martina Stein who typed parts of this paper with considerable technical expertise. The work of the authors was supported by the Sonderforschungsbereich 475, Komplexitätsreduktion in multivariaten Datenstrukturen (Teilprojekt A2) and in part by a NIH grant award IR01GM072876:01A1 and a BMBF-grant SKAVOE.

References

Atkinson, A. (1970). The design of experiments to estimate the slope of a response surface. *Biometrika*, 57:319–328.

Dette, H. (1995). Designing of experiments with respect to “standardized” optimality criteria. *Journal of the Royal Statistical Society, Ser. B*, 59:97–110.

Dette, H. and Braess, D. (2007). On the number of support points of maximin and bayesian d-optimal designs in nonlinear regression models. *Annals of Statistics*, 35:772–792.

Dette, H., Melas, V., and Pepelyshev, A. (2003). Standardized maximin e-optimal designs for the michaelis-menten model. *Statistica Sinica*, 13:1147–1163.

Dette, H., Melas, V., and Pepelyshev, A. (2004). Optimal designs for a class of nonlinear regression models. *Ann. Statist.*, 32:2142–2167.

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

Fedorov, V. and Müller, W. (1997). Another view on optimal design for estimating the point of extremum in quadratic regression. *Metrika*, 46:147–157.

Hader, R. and Park, S. (1978). Slope-rotatable central composite designs. *Technometrics*, 20:413–417.

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

Kiefer, J. (1974). General equivalence theory for optimum designs (approximate theory). *Ann. Statist.*, 2:849–879.

Lau, T.-S. and Studden, W. (1985). Optimal designs for trigonometric and polynomial regression using canonical moments. *Ann. Statist.*, 13:383–394.

Lestrel, P. (1997). *Fourier Descriptors and Their Applications in Biology.* Cambridge University Press, New York.

Mandal, N. and Heiligers, B. (1992). Minimax designs for estimating the optimum point in a quadratic response surface. *Journal of Statistical Planning and Inference*, 31:235–244.

Melas, V., Pepelyshev, A., and Cheng, R. (2003). Designs for estimating an extremal point of quadratic regression models in a hyperball. *Metrika*, 58:193–208.

Mukerjee, R. and Huda, S. (1985). Minimax second- and third-order designs to estimate the slope of a response surface. *Biometrika*, 72:173–178.

Müller, C. (1995). Maximin efficient designs for estimating nonlinear aspects in linear models. *Journal of Statistical Planning and Inference*, 44:117–132.

Müller, C. and Pázman, A. (1998). Applications of necessary and sufficient conditions for maximum efficient designs. *Metrika*, 48:1–19.

Murthy, V. and Studden, W. (1972). Optimal designs for estimating the slope of a polynomial regression. *J. Am. Statist. Assoc.*, 67:869–873.

Myres, R. and Lahoda, S. (1975). A generalization of the response surface mean square error criterion with a specific application to the slope. *Technometrics*, 17:481–486.

Ott, L. and Mendenhall, W. (1972). Designs for estimating the slope of a second order linear model. *Technometrics*, 14:341–353.

Pronzato, L. and Walter, E. (1993). Experimental design for estimating the optimum point in a response surface. *Acta Applicandae Mathematicae*, 33:45–68.

Pukelsheim, F. (1993). *Optimal Design of Experiments*. John Wiley & Sons, New York.

Pukelsheim, F. and Rieder, S. (1992). Efficient rounding of approximate designs. *Biometrika*, 79:763–770.

Sahm, M. (1998). *Optimal designs for estimating individual coefficients in polynomial regression.* PhD thesis, Fakultät für Mathematik, Ruhr-Universität Bochum, Germany.

Silvey, S. (1980). *Optimal Design.* Chapman and Hall, London.

Szegö, G. (1975). *Orthogonal Polynomials.* American Mathematical Society, Providence, R.I.

Wu, H. (2002). Optimal designs for first-order trigonometric regression on a partial circle. *Statist. Sinica*, 12:917–930.

Zen, M.-M. and Tsai, M.-H. (2004). Criterion-robust optimal designs for model discrimination and parameter estimation in fourier regression models. *J. Stat. Plann. Inference.*, 124:475–487.