A functional-algebraic determination of D-optimal designs for trigonometric regression models on a partial circle

Holger Dette
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany
email: holger.dette@ruhr-uni-bochum.de
FAX: +49 2 34 70 94 559

Viatcheslav B. Melas
St. Petersburg State University
Department of Mathematics
St. Petersburg
Russia
email: v.melas@pobox.spbu.ru

Stefanie Biedermann
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany
email: stefanie.biedermann@ruhr-uni-bochum.de

March 29, 2001

Abstract

We investigate the D-optimal design problem in the common trigonometric regression model, where the design space is a partial circle. The task of maximizing the criterion function is transformed into the problem of determining an eigenvalue of a certain matrix via a differential equation approach. Since this eigenvalue is an analytic function of the length of the design space, we can make use of a Taylor expansion to provide a recursive algorithm for its calculation. Finally, this enables us to determine Taylor expansions for the support points of the D-optimal design.

AMS Subject Classification: 62K05
Keywords and Phrases: trigonometric regression, D-optimality, implicit function theorem, differential equation.

1 Introduction

Consider the common trigonometric regession model

$$y = \beta_0 + \sum_{j=1}^{m} \beta_{2j-1} \sin(jt) + \sum_{j=1}^{m} \beta_{2j} \cos(jt) + \varepsilon, \quad t \in [c,d], \tag{1.1}$$

$-\infty < c < d < \infty$, which is widely used to describe a periodic relation between some observations and the points, where these observations are taken. We will assume that the errors in model (1.1) are i.i.d. random values with zero expectation and a finite variance. Much effort has been devoted to the problem of designing experiments for model (1.1) to obtain efficient estimates for the coefficients. Most authors concentrate on the design space $[-\pi, \pi]$, but Hill (1978) and Kitsos, Titterington and Torsney (1988) point out that in many applications it is impossible to take observations on the full circle $[-\pi, \pi]$. We refer to Kitsos, Titterington and Torsney (1988) for a concrete example, who investigated a design problem in rhythmometry involving circadian rhythm exhibited by peak expiratory flow, for which the design region has to be restricted to a partial cycle of the complete 24-hour period.

In the present paper, we address the question of designing experiments in trigonometric models, where the design space is not necessarily the full circle but an arbitrary interval $[c, d] \subset \mathbb{R}$. Recently, Dette, Melas and Pepelyshev (2001) considered D-optimal designs for estimating the coefficients in this model using a functional approach. In the present paper, we combine this with an algebraic approach, which is developed by transforming the original problem into a differential equation problem leading us to an eigensystem of a certain matrix. Taking into account some well-known facts about matrix-algebra, we obtain Taylor expansions for the sought eigenvalue and the corresponding (normalized) eigenvector. Finally, these findings yield a method to calculate Taylor expansions for the support points of the D-optimal design.

2 Preliminary results for D-optimal designs in trigonometric regression models on a partial circle

Consider the trigonometric regression model (1.1), define $\beta = (\beta_0, \beta_1, \ldots, \beta_{2m})^T$ as the vector of parameters and

$$f(t) = (1, \sin t, \cos t, \ldots, \sin(mt), \cos(mt))^T = (f_0(t), \ldots, f_{2m}(t))^T \tag{2.1}$$

as the vector of regression functions. Due to the 2π-periodicity of the regression functions we restrict ourselves without loss of generality to design spaces with length $d - c \le 2\pi$. It was shown in Dette, Melas and Pepelyshev (2001), that a D-optimal design on the interval $[c, d]$, which means a probability measure ξ on $[c, d]$ maximizing the determinant of the information matrix

$$M(\xi) = \int f(t) f^T(t) d\xi(t) \in \mathbb{R}^{2m+1 \times 2m+1} \ , \tag{2.2}$$

can be obtained from a D-optimal design $\xi(a)$ on the symmetric interval $[-a, a]$ by subtracting the value $(d - c)/2$ from the support points of $\xi(a)$. Therefore, it is sufficient to study D-optimal designs on the interval $[-a, a]$, $0 < a \le \pi$. Moreover, it was proved in the same reference that for

$$a \ge \hat{a} := \pi(1 - \frac{1}{2m+1}) \tag{2.3}$$

designs with information matrix $M(\xi) = diag\{1, 1/2, 1/2, \ldots, 1/2\}$ are D-optimal, such as, in particular, the design

$$\xi^* = \begin{pmatrix} t_1^* & \cdots & t_{2m+1}^* \\ \frac{1}{2m+1} & \cdots & \frac{1}{2m+1} \end{pmatrix},$$

where $t_i^* = 2\pi(i - 1 - m)/(2m + 1)$, $i = 1, \ldots, 2m + 1$. The case $a < \pi(1 - 1/(2m + 1))$ was studied by means of a functional approach.

In the present paper, we will obtain a number of new results concerning D-optimal designs on the interval $[-a, a]$, $0 < a < \pi(1 - 1/(2m + 1))$. In particular, these results provide a more efficient version of the functional approach for solving the problem at hand.

3 The differential equation and the eigenvalue problem

To proceed to the differential equation form of the problem, we will need the following auxiliary result. The proof can be found in Dette, Melas and Pepelyshev (2001).

Lemma 3.1. *a) There exists a unique D-optimal design for the trigonometric regression model on the segment $[-a,a]$, $0 < a < \pi(1-1/(2m+1))$, which is of the form*

$$\xi = \xi(a) \;=\; \begin{pmatrix} -t_m & \dots & -t_1 & t_0 & t_1 & \dots & t_m \\ \frac{1}{2m+1} & \dots & \frac{1}{2m+1} & \frac{1}{2m+1} & \frac{1}{2m+1} & \dots & \frac{1}{2m+1} \end{pmatrix}, \tag{3.1}$$

$t_m = a,\ t_0 = 0.$

b) For any design ξ of the form (3.1) we have

$$\begin{aligned} \det M(\xi) &= C\ \phi(x,a) \\ &= \frac{2^{2m^2}}{(2m+1)^{2m+1}} \prod_{i=1}^{m} (1-x_i^2)(1-x_i)^2 \prod_{1\le i<j\le m} (x_j - x_i)^4, \end{aligned} \tag{3.2}$$

where $x = (x_1,\dots,x_m)$, $x_i = \cos(t_i)$, $i = 1,\dots,m$. Moreover, for fixed a the function $\phi(x,a)$ is strictly concave.

Note that due to formula (3.2), the support points of the D-optimal design on the interval $[-a,a]$, $t_i(a)$, $i = 1,\dots,m$, can be written in the form

$$t_i(a) = \arccos(x_i^*),$$

where $x^* = (x_1^*,\dots,x_m^*)$ is the unique point of maximum of the function $\phi(x,a)$ on the set

$$\mathcal{X} = \{x = (x_1,\dots,x_m);\ 0 < x_1 < \ldots < x_m = \cos(a)\}.$$

Calculating the first partial derivatives of $\phi(x,a)$, we obtain

$$\frac{1}{1+x_i} - \frac{3}{1-x_i} + \frac{4}{x_i - 1 + \alpha} + \sum_{j=1, j\neq i}^{m-1} \frac{4}{x_i - x_j} = 0,$$

$i = 1,\dots,m-1$ with $x_i = x_i^*$, where $\alpha = 1-\cos(a)$. Consider the supporting polynomial

$$\psi(z) = \prod_{i=1}^{m-1} (z - x_i^*) = z^{m-1} + \sum_{i=0}^{m-2} \psi_i z^i.$$

Applying the following well-known equality (see, for instance, Fedorov (1972)),

$$\sum_{j=1, j\neq i}^{m-1} \frac{1}{x_i^* - x_j^*} = \frac{1}{2}\,\frac{\psi''(x_i^*)}{\psi'(x_i^*)}, \quad i = 1,\dots,m-1,$$

we receive the relations

$$\frac{-1-2z}{1-z^2} + \frac{2}{z-1+\alpha} + \frac{\psi''(z)}{\psi'(z)} = 0$$

for $z = x_1^*, \dots, x_{m-1}^*$. Multiplying the equation by the common denominator, we obtain

$$(1 - z^2)(z - 1 + \alpha)\psi''(z) + (-4z^2 + (1 - 2\alpha)z + 3 - \alpha)\psi'(z) = 0$$

again for $z = x_1^*, \dots, x_{m-1}^*$.

Since on the left hand side, there is a polynomial of degree m vanishing at the $m-1$ points $x_i^*,\ i = 1, \dots, m-1$, we can equate this to the polynomial ψ multiplied by a linear factor, so that the problem turns out to be one of solving a second order differential equation

$$\begin{aligned} P(z) &:= (1 - z^2)(z - 1 + \alpha)\psi''(z) + (-4z^2 + (1 - 2\alpha)z + 3 - \alpha)\psi'(z) \\ &- (\vartheta_0 z + \lambda)\psi(z) \equiv 0 \end{aligned} \tag{3.3}$$

where $\vartheta_0 = -(m-1)(m+2)$ is obtained by comparing coefficients of z^m and λ is an unknown real constant.

Using that the solution ψ^* of the differential equation is supposed to be a polynomial of degree $m-1$, we can rewrite $P(z)$ in the matrix-vector form

$$P(z) = (z^m, \dots, z, 1)\ A(\lambda, \alpha)\ \psi, \tag{3.4}$$

where $\psi = (\psi_{m-1}, \dots, \psi_0)^T$ and $A = A(\lambda, \alpha)$ is some $(m+1) \times m$-matrix. Note that the first row of A consists of zeros. Let $B = B(\lambda, \alpha)$ be the matrix obtained from A by deleting the first row with elements $b_{i,j} = (B(\lambda, \alpha))_{i,j},\ i, j = 1, \dots, m$. Comparing the coefficients of the monomials $z^j,\ j = 0, \dots, m$ in (3.4) yields

$$b_{i,j} = \begin{cases} -(m-j)(m-j+3) - \vartheta_0 & j - i = 1 \\ (m-j)((1-\alpha)(m-j+1) - 1) - \lambda & j - i = 0 \\ (m-j)(m-j-\alpha+2) & j - i = -1 \\ (m-j)(m-j-1)(\alpha-1) & j - i = -2 \\ 0 & otherwise \end{cases}. \tag{3.5}$$

Note that the matrix B is of the form $B = B(\lambda, \alpha) = \tilde{B}(\alpha) - \lambda\ I_m$, and λ is an eigenvalue of the matrix $\tilde{B}(\alpha)$. Therefore, we can rewrite equation (3.3) in the form

$$(\tilde{B}(\alpha) - \lambda\ I_m)\ \psi = 0. \tag{3.6}$$

For known λ, we conclude from (3.5) that the vector ψ can be calculated by the following recursive relations

$$\begin{aligned} \psi_{m-1} &= 1 \\ \psi_\nu &= -\sum_{j=\nu+1}^{m-1} b_{m-\nu-1, m-j}\ \psi_j\ /\ b_{m-\nu-1, m-\nu} \end{aligned} \tag{3.7}$$

$\nu = m-2, m-3, \dots, 0$.

A method to calculate the eigenvalue of interest will be described in the following section. Our approach based on the algebraic equation (3.6) will be called an algebraic approach. Note that a similar method was suggested in Dette, Haines and Imhof (1999) and in Melas (1999) for studying (locally) D-optimal designs for rational models. In the present paper, we will combine this approach with the functional approach suggested in Dette, Melas and Pepelyshev (2001).

4 A functional-algebraic approach

Consider the function

$$g(\lambda,\alpha) = \det(\tilde{B}(\alpha) - \lambda\ I_m).$$

The unknown value λ in equation (3.3) is a function of α ($\lambda^*(\alpha)$, say), to be explicitly given by the equation

$$g(\lambda,\alpha) = 0.$$

Since λ is a simple eigenvalue of $\tilde{B}(\alpha)$ (recursive formula (3.7) shows that the corresponding normalized eigenvector is unique), the following equation holds.

$$\frac{d}{d\lambda} g(\lambda,\alpha)\Big|_{\lambda=\lambda^*(\alpha)} \neq 0$$

Due to the implicit function theorem (see Gunning and Rossi (1965)), $\lambda^*(\alpha)$ is a real analytic function on the interval $(0,\hat{\alpha})$, where $\hat{\alpha} = 1-\cos(\hat{a})$ and $\hat{a}$ is defined in (2.3). This also follows from the fact that simple eigenvalues of a matrix are real analytic (see Lancaster (1969)). Consequently, the function $\lambda^*(\alpha)$ can be expanded into a Taylor series on this interval. To expand this function in a neighbourhood of the origin, we must continue it to the interval $(-\hat{\alpha},\hat{\alpha})$. So our aim is to find the limit of $\lambda^*(\alpha)$ when $\alpha \to 0$, which can be realized by taking the limit in (3.3). Since all the points in the D-optimal design tend to zero, it follows that $x_i^* \to 1$, $i = 1,\dots,m-1$ and for the supporting polynomial $\psi(z) \to (1-z)^{m-1}$. By direct calculations, we obtain

$$\lim_{\alpha\to 0} \lambda^*(\alpha) = 1 - m^2.$$

Hence the function

$$\hat{\lambda}(\alpha) = \begin{cases} \lambda^*(\alpha) & 0 < \alpha < \hat{\alpha} \\ \lambda^*(-\alpha) & 0 > \alpha > -\hat{\alpha} \\ 1-m^2 & \alpha = 0 \end{cases}$$

is real analytic on the interval $(-\hat{\alpha},\hat{\alpha})$. Consider its Taylor expansion

$$\hat{\lambda}(\alpha) = \sum_{i=0}^{\infty} \lambda_{(i)}\ \alpha^i, \quad \lambda_{(0)} = 1 - m^2, \tag{4.1}$$

and let

$$\begin{aligned} \lambda_{<n>}(\alpha) &= \sum_{i=0}^{n} \lambda_{(i)}\ \alpha^i, \\ \left(g(\lambda_{<n>}(\alpha),\alpha)\right)_{(n)} &= \frac{1}{n!}\frac{\partial^n}{\partial\alpha^n}\ g(\lambda_{<n>}(\alpha),\alpha)\Big|_{\alpha=0}. \end{aligned}$$

To determine the coefficients $\lambda_{(i)}$ in this expansion, we will use the following recursive formulas, which have been explicitly found in Dette, Melas and Pepelyshev (2000):

$$\begin{aligned} \lambda_{(n+1)} &= -J^{-1}(0)\left(g(\lambda_{<n>}(\alpha),\alpha)\right)_{(n+1)}, \quad n = 0,1,\dots \\ J(\alpha) &= \frac{\partial}{\partial\lambda}\ g(\lambda,\alpha). \end{aligned}$$

The first values of the scaled coefficients $\bar{\lambda}_{(i)} = \lambda_{(i)} 2^i$ are given in Table 1.

Table 1: *Coefficients* $\bar{\lambda}_{(i)} = 2^i \lambda_{(i)}$ *in the expansion (4.1) of the eigenvalue and coefficients* $\bar{\psi}_{j(i)}$ *in the expansion (4.2) of the components of the corresponding eigenvector* $(\hat{\psi}_0, \ldots, \hat{\psi}_{m-1})$.

	i	0	1	2	3	4	5
$m=2$	$\bar{\psi}_{0(i)}$	-1	.85714	-.06997	-.02856	-.00886	-.00133
	$\bar{\lambda}_{(i)}$	-3	-.57143	-.27988	-.11424	-.03544	-.00533
$m=3$	$\bar{\psi}_{0(i)}$	1	-1.81818	.67618	-.02666	-.00204	.00190
	$\bar{\psi}_{1(i)}$	-2	1.81818	-.07012	-.03345	-.01963	-.01248
	$\bar{\lambda}_{(i)}$	-8	-1.09091	-.42074	-.20068	-.11779	-.07489
$m=4$	$\bar{\psi}_{0(i)}$	-1	2.80000	-2.22277	.48317	-.00808	.00222
	$\bar{\psi}_{1(i)}$	3	-5.00000	2.29169	-.05781	-.01309	-.00261
	$\bar{\psi}_{2(i)}$	-3	2.80000	-.06892	-.03375	-.02060	-.01400
	$\bar{\lambda}_{(i)}$	-15	-1.60000	-.55138	-.27001	-.16480	-.11199
$m=5$	$\bar{\psi}_{0(i)}$	1	-3.78947	4.74901	-2.23711	.32001	-.00380
	$\bar{\psi}_{1(i)}$	-4	11.36842	-9.56600	2.36016	-.03433	-.00019
	$\bar{\psi}_{2(i)}$	6	-11.36842	4.88488	-.08948	-.02490	-.00831
	$\bar{\psi}_{3(i)}$	-4	3.78947	-.06792	-.03357	-.02072	-.01430
	$\bar{\lambda}_{(i)}$	-24	-2.10526	-.67923	-.33566	-.20720	-.14296

Note that since the eigenvectors of a matrix are real analytic functions (see Lancaster (1969)), the coefficients $\psi_j = \psi_j(\alpha)$, $j = m-2, \ldots, 0$ are real analytic functions on the interval $(0, \hat{\alpha})$. So the problem of determining the components of the (normalized) eigenvector can be dealt with analogously to that of calculating the eigenvalue. By the relations

$$\hat{\psi}_j\ (\alpha) = \begin{cases} \psi_j\ (\alpha) & 0 < \alpha < \hat{\alpha} \\ \psi_j\ (-\alpha) & 0 > \alpha > -\hat{\alpha} \\ \psi_j\ (0) & \alpha = 0 \end{cases}$$

where $\psi_j(0) = (-1)^{m-j-1}(m-1)!\ /\ (j\ !(m-j-1)!)$ these functions can be analytically expanded on the interval $(-\hat{\alpha}, \hat{\alpha})$. The Taylor expansions

$$\hat{\psi}_j(\alpha) = \sum_{i=0}^{\infty} \bar{\psi}_{j(i)}\ \alpha^i /\ 2^i \tag{4.2}$$

can be constructed using the recursive formulas (3.5). The first coefficients are listed in Table 1.

Using the values of the $\psi_i(\alpha)$ for the components of the eigenvector, the Taylor expansions of the functions (which give the support points of the D-optimal design) $t_i(a)$, $i = 1, \ldots, m-1$ can be constructed as follows. Note that these functions are real analytic because the roots of a polynomial are real analytic functions of its coefficients.

Let us define the polynomial $\rho(u, \alpha)$ by the relation

$$\rho(u, \alpha) = \alpha^{1-m}\ \psi(1 - \alpha u).$$

Denote by $u_i(0)$, $i = 1, \ldots, m-1$, the roots of $\rho(u, 0) = const\ P_{m-1}^{(1,1/2)}(2u-1)$, where $P_{m-1}^{(\beta,\gamma)}$ is the Jacobi polynomial with parameters (β, γ) of degree $m-1$. Construct expansions of the solutions $u_i(\alpha) = u(\alpha)$ of the equation $\rho(u,\alpha) = 0$ with the initial condition $u(0) = u_i(0)$ by the functional approach described above and return to the original variables $t_i(a) = \arccos(x_i(\alpha)) = \arccos(1 - \alpha\ u_i(\alpha)),\ \alpha = 1 - \cos(a),\ i = 1, \ldots, m-1$.

Proceeding as described above, we obtained the first coefficients of the Taylor expansions for the support points $t_i(a),\ i = 1, \ldots, m-1$ of the D-optimal design for the trigonometric regression model (1.1) on the interval $[-a, a]$ (if $a < \hat{a}$), which are the same as in Dette, Melas and Pepelyshev (2001). However, the present approach appears to be preferable in computer time and memory compared to the direct functional approach of that paper. This is not surprising, since the algebraic-analytical approach takes into account the special structure of the problem at hand.

Acknowledgements. The work of V.B. Melas was partially supported by the Russian Foundation of Basic Research (grant No. 00-01-00495). The work of S. Biedermann and H. Dette was supported by the Deutsche Forschungsgemeinschaft (SFB 475: Komplexitätsreduktion in multivariaten Datenstrukturen). Parts of this paper were written during a visit of H. Dette at Purdue University and this author would like to thank the Department of Statistics for its hospitality. The authors are also grateful to A. Pepelyshev for his help in computer calculations.

References

H. Dette, L. Haines, L. Imhof (1999). Optimal designs for rational models and weighted polynomial regression. The Annals of Statistics, Vol. 27. N 4, 1272-1293.

H. Dette, V.B. Melas, A. Pepelyshev (2000). Optimal designs for estimating individual coefficients - a functional approach. Preprint Ruhr-Universtät Bochum.
http://www.ruhr-uni-bochum.de/mathematik3/preprint.htm.

H. Dette, V.B. Melas, A. Pepelyshev (2001). D-optimal designs for trigonometric regression models on a partial circle. Preprint, Ruhr-Universität Bochum.
http://www.ruhr-uni-bochum.de/mathematik3/preprint.htm.

H. Dette, V.B. Melas, (2000). Optimal designs for estimating individual coefficients in Fourier regression models. Preprint, Ruhr-Universität Bochum.
http://www.ruhr-uni-bochum.de/mathematik3/preprint.htm.

V.V. Fedorov (1972). Theory of optimal experiments. Academic Press, New York.

R.C. Gunning, H. Rossi (1965). Analytical Functions of Several Complex Variables. Prentice Hall, New York.

P.D.H. Hill (1978). A note on the equivalence of D-optimal design measures for three rival linear models. Biometrika 65, 666-667.

C.P. Kitsos, D.M. Titterington, B. Torsney (1988). An optimal design problem in rhythmometry. Biometrics 44, 657-671.

P. Lancaster (1969). Theory of matrices. Academic Press, New York.

V.B. Melas (1978). Optimal designs for exponential regression. Math. Oper. Forsch. Stat., Ser. Statistics, 9, 45-59.

V.B. Melas (1999). Locally optimal designs of experiments. Publishers of St. Petersburg State Technical University, St. Petersburg (Russian).

F. Pukelsheim (1993). Optimal Design of Experiments. Wiley, New York.

S.D. Silvey (1980). Optimal design. Chapman and Hall, London.