Statistica Sinica **19** (2009), 1587-1601

OPTIMAL DESIGNS FOR ESTIMATING PAIRS OF COEFFICIENTS IN FOURIER REGRESSION MODELS

Holger Dette[1], Viatcheslav B. Melas[2] and Petr Shpilev[2]

[1]*Ruhr-Universität Bochum and* [2]*St. Petersburg State University*

Abstract: In the common Fourier regression model we investigate the optimal design problem for estimating pairs of the coefficients, where the explanatory variable varies in the interval $[-\pi, \pi]$. L-optimal designs are considered and for many important cases L-optimal designs can be found explicitly, where the complexity of the solution depends on the degree of the trigonometric regression model and the order of the terms for which the pair of the coefficients has to be estimated.

Key words and phrases: Equivalence theorem, Fourier regression models, L-optimal designs, parameter subsets.

1. Introduction

Fourier regression models of the form

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

are widely used in applications to describe periodic phenomena. Typical subject areas include engineering (see e.g., McCool (1979)), medicine (see e.g., Kitsos, Titterington and Torsney (1988) and biology (see the collection of research papers edited by Lestrel (1997)). Applications of trigonometric regression models also appear in two dimensional shape analysis (see e.g., Young and Ehrlich (1977) and Currie, Ganeshanandam, Noition, Garrick, Shelbourne and Oragguzie (2000) among many others). An optimal design can improve the efficiency of the statistical analysis substantially, and the problem of designing experiments for Fourier regression models has been discussed by several authors. Optimal designs with respect to Kiefer's ϕ_p-optimality criteria have been studied by Karlin and Studden (1966) and Wu (2002) among others (see also (Pukelsheim, 1993, p.241)), while Lau and Studden (1985) discuss the problem of constructing robust designs if the degree m in model (1.1) is not known. Designs for identifying the degree m have been determined by Biedermann, Dette and Hoffmann (2007), Dette and Haller (1998), and Zen and Tsai (2004). More recent work discussed the problem of constructing optimal designs for the estimation of a particular coefficient in the Fourier regression model (1.1) (see Dette and Melas (2002, 2003), Dette, Melas and Pepelysheff (2002), and Dette, Melas and Shpilev (2007)). These authors also demonstrated that uniform designs are optimal for estimating a subset of the coefficients $\{\beta_{2i_1-1}, \beta_{2i_1}, \ldots, \beta_{2i_r-1}, \beta_{2i_r}\}$, where $1 \le i_1 < \ldots < i_r \le m$, $r \in \{1, \ldots, m\}$.

The main purpose of the present paper is to obtain further insight into the construction of optimal designs for estimating parameter subsystems in the Fourier regression model (1.1). In particular we are interested in the L-optimal design problem for estimating pairs of the coefficients $\{\beta_{2i_1}, \beta_{2i_2}\}$ and $\{\beta_{2j_1-1}, \beta_{2j_2-1}\}$, where $i_1, i_2 \in \{0, \ldots, m\}$, $j_1, j_2 \in \{1, \ldots, m\}$. The precise estimation of specific pairs of coefficients is of particular interest because in many biological applications, such as two dimensional shape analysis, one or two coefficients have a concrete biological interpretation (see e.g., Young and Ehrlich (1977), Currie et al. (2000)). In Section 2, we introduce the general notation and state several preliminary results. We formulate and prove a particular version of the equivalence theorem for L-optimal designs, an important tool for the determination of the optimal designs, in Section 3. Here L-optimal designs are found explicitly for several cases. Finally, several examples are presented in Section 4 to illustrate the theoretical results.

2. *L*-optimal Designs

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

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

as the vector of regression functions. Following Kiefer (1974) we call any probability measure ξ on the design space $[-\pi, \pi]$ with finite support an (approximate) design. The support points of the design ξ give the locations where observations are taken, while the weights give the corresponding proportions of the total number of observations to be taken at these points. If the design ξ puts masses ξ_i at the points t_i $(i = 1, \ldots, k)$ and n uncorrelated observations can be taken, then the quantities $\xi_i n$ are rounded to integers such that $\sum_{i=1}^k n_i = n$ and the experimenter takes n_i observations at each t_i $(i = 1, \ldots, k)$. In this case the covariance matrix of the least squares estimate for the parameter β in the trigonometric regression model (1.1) is approximately given by $(\sigma^2/n)M^{-1}(\xi)$, where

$$M(\xi) = \left(\int_{-\pi}^{\pi} f(t) f^T(t) d\xi(t) \right) \in R^{(2m+1)\times(2m+1)} \tag{2.1}$$

denotes the information matrix of the design ξ (see Pukelsheim (1993)). Note that for a symmetric design ξ, after an appropriate permutation $P \in R^{(2m+1)\times(2m+1)}$

of the order of the regression functions, the information matrix (2.1) is block diagonal, that is

$$\widetilde{M}(\xi) = PM(\xi)P = \begin{pmatrix} M_c(\xi) & 0 \\ 0 & M_s(\xi) \end{pmatrix} \tag{2.2}$$

with blocks given by

$$\begin{aligned} M_c(\xi) &= \left(\int_{-\pi}^{\pi} \cos(it)\cos(jt) d\xi(t) \right)_{i,j=0}^{m}, \\ M_s(\xi) &= \left(\int_{-\pi}^{\pi} \sin(it)\sin(jt) d\xi(t) \right)_{i,j=1}^{m}. \end{aligned} \tag{2.3}$$

For a given matrix

$$L = \sum_{i=0}^{k} l_i l_i^T, \quad k \le 2m \tag{2.4}$$

with vectors $l_i \in \mathbb{R}^{2m+1}$, the class Ξ_L is defined as the set of all approximate designs for which the linear combinations of the parameters $l_i^T\beta$, $i = 0, \ldots, k$ are estimable, that is $l_i \in \text{Range}(M(\xi))$; $i = 0, \ldots, k$. Similarly, the sets Ξ_s and Ξ_c are defined as the sets of all designs for which the matrices $M_s(\xi)$ and $M_c(\xi)$ are nonsingular, respectively. Finally a design ξ^* is called L-optimal if $\xi^* = \arg\min_{\xi\in\Xi_L} trLM^+(\xi)$, where L is a fixed and nonnegative definite matrix, and for a given matrix A the matrix A^+ is the pseudo-inverse (see Rao (1968)). The following result gives a characterization of L-optimal designs that is particularly useful for designs with a singular information matrix. For a nonsingular information matrix this theorem was formulated and proved in Ermakov and Zhigljavsky (1987). The result is stated for a general regression model $y = \beta^T f(t) + \varepsilon$ with $2m+1$ regression functions and a general design space χ.

Theorem 2.1. *Let $L \in \mathbb{R}^{(2m+1)\times(2m+1)}$ denote a given and nonnegative definite matrix of the form* (2.4), *and assume that the set of information matrices is compact.*

(1) *The design ξ is an element of the class Ξ_L if and only if $l_i^T M^+(\xi)M(\xi) = l_i^T$, $i = 0, \ldots, k$.*

(2) *The design ξ is L-optimal if*

$$\max_{t\in\chi} \varphi(t, \xi^*) = trLM^+(\xi^*), \tag{2.5}$$

where $\varphi(t, \xi) = f^T(t)M^+(\xi)LM^+(\xi)f(t)$. Moreover, the equality

$$\varphi(t_i, \xi^*) = trLM^+(\xi^*) \tag{2.6}$$

holds for any $t_i \in supp(\xi^*)$.

Proof. The first part of this theorem was proved in Rao (1968). For a proof of the second part we use the following lemma.

Lemma 2.1. *If A is a nonnegative definite matrix, then $A + A^+ \geq 2A^+A$ with equality if and only only if $A = A^+A$.*

Proof of Lemma 2.1. Let $A = B^TB$, where $B \in \mathbb{R}^{n\times n}$. By definition of the pseudo-inverse we have $A^+ = B^+B^{+T}$, and so

$$\begin{aligned} A + A^+ - 2A^+A &= B^TB + B^+B^{+T} - 2B^+B^{+T}B^TB \\ &= (B^T - B^+)(B - B^{+T}) = (B - B^{+T})^T(B - B^{+T}) \geq 0, \end{aligned}$$

where the second equality is obtained from the identity $B^+B^{+T}B^TB = B^+(BB^+)^TB = B^+BB^+B = B^+B$.

Now we return to the proof of the second part of Theorem 2.1. We begin by showing that a L-optimal design satisfies (2.5) and (2.6). For any $\alpha \in (0,1)$, define $\xi_\alpha = (1-\alpha)\xi^* + \alpha\xi_t$, where ξ_t denotes the Dirac measure at the point t, and assume that the design ξ^* is L-optimal. In this case the inequality $trLM^+(\xi_\alpha) \geq trLM^+(\xi^*)$ is satisfied for all $\alpha \in (0,1)$, which implies $(\partial trLM^+(\xi_\alpha))/(\partial\alpha)\ |_{\alpha=0+} \geq 0$. On the other hand we obtain, observing the identity

$$\frac{\partial LM^+(\xi_\alpha)M(\xi_\alpha)}{\partial\alpha} = 0 = L\frac{\partial M^+(\xi_\alpha)}{\partial\alpha}M(\xi_\alpha) + LM^+(\xi_\alpha)\frac{\partial M(\xi_\alpha)}{\partial\alpha},$$

the inequality

$$\begin{aligned} \frac{\partial trLM^+(\xi_\alpha)}{\partial\alpha}\Big|_{\alpha=0+} &= tr\Big\{L\frac{\partial M^+(\xi_\alpha)}{\partial\alpha}\Big\}\Big|_{\alpha=0+} \\ &= tr\Big\{-LM^+(\xi_\alpha)\frac{\partial M(\xi_\alpha)}{\partial\alpha}M^+(\xi_\alpha)\Big\}\Big|_{\alpha=0+} \\ &= trLM^+(\xi^*) - trLM^+(\xi^*)M(\xi_t)M^+(\xi^*) \\ &= trLM^+(\xi^*) - \varphi(t,\xi^*) \geq 0. \end{aligned}$$

Therefore we have $\varphi(t,\xi^*) \leq trLM^+(\xi^*)$ for all t. Moreover, the equality

$$\int \varphi(t,\xi)\xi(dt) = \int trLM^+(\xi)f(t)f^T(t)M^+(\xi)\xi(dt) = trLM^+(\xi)$$

is obviously fulfilled for any design $\xi \in \Xi_L$. It now follows that (2.5) and (2.6) are satisfied.

In general an analytical determination of L-optimal designs is very difficult. Theorem 2.1 can be used to check the optimality of a given design numerically. We use this characterization in the following section.

3. Analytical Solutions of the *L*-optimal Design Problem

In the present section we develop explicit solutions of the L-optimal design problem in the Fourier regression model (1.1) for several parameter subsystems. We begin with the problem of constructing optimal designs for estimating two coefficients corresponding to the sinus- or cosinus functions. More precisely, define $e_j \in \mathbb{R}^{2m+1}$ as the jth unit vector and consider the matrices

$$L_{(2\lfloor m/2\rfloor-1,\ 4\lfloor m/2\rfloor-1)} = e_{2\lfloor m/2\rfloor-1}e^T_{2\lfloor m/2\rfloor-1} + e_{4\lfloor m/2\rfloor-1}e^T_{4\lfloor m/2\rfloor-1}$$
$$L_{(2\lfloor m/2\rfloor,\ 4\lfloor m/2\rfloor)} = e_{2\lfloor m/2\rfloor}e^T_{2\lfloor m/2\rfloor} + e_{4\lfloor m/2\rfloor}e^T_{4\lfloor m/2\rfloor}.$$

We call an L-optimal design with the matrix $L_{(2\lfloor m/2\rfloor-1,\ 4\lfloor m/2\rfloor-1)}$ or $L_{(2\lfloor m/2\rfloor,\ 4\lfloor m/2\rfloor)}$ an L-optimal design for estimating the pair of coefficients $\beta_{2\lfloor m2\rfloor-1}$, $\beta_{4\lfloor m/2\rfloor-1}$ or $\beta_{2\lfloor m/2\rfloor}$, $\beta_{4\lfloor m/2\rfloor}$, respectively.

Theorem 3.1. *Consider the trigonometric regression model* (1.1) *with* $m > 3$.

(1) *The design*

$$\xi^*_{(2\lfloor m/2\rfloor-1,\ 4\lfloor m/2\rfloor-1)} = \begin{pmatrix} -t_n & -t_{n-1} & \dots & -t_1 & t_1 & \dots & t_n \\ \frac{1}{2n} & \frac{1}{2n} & \dots & \frac{1}{2n} & \frac{1}{2n} & \dots & \frac{1}{2n} \end{pmatrix},$$

with $n = 2\lfloor m/2\rfloor$, $t_i = 2\lfloor i/2\rfloor(\pi/n) + (-1)^{(i-1)}x$, $x = (2\arctan(\sqrt[4]{5}))/n$ *is L-optimal for estimating the pair of coefficients* $\beta_{2\lfloor m/2\rfloor-1}$, $\beta_{4\lfloor m/2\rfloor-1}$. *Moreover,* $trLM^+(\xi^*_{(2\lfloor m/2\rfloor-1,\ 4\lfloor m/2\rfloor-1)}) = (\sqrt{5}/2) + (3/2)$.

(2) *For any* $\alpha \in [0, \omega_n]$ *the design*

$$\xi^*_{(2\lfloor m/2\rfloor,\ 4\lfloor m/2\rfloor)} = \begin{pmatrix} -\pi & -t_{n-1} & \dots & -t_1 & 0 & t_1 & \dots & t_{n-1} & \pi \\ \omega_n - \alpha & \omega_{n-1} & \dots & \omega_1 & \omega_0 & \omega_1 & \dots & \omega_{n-1} & \alpha \end{pmatrix}$$

with $n = 2\lfloor m/2\rfloor$, $t_i = ((i-1)\pi)/n$, $i = 2, \dots, n$, $\omega_0 = \sqrt{5}\omega_1$, $\omega_1 = (\sqrt{5}-1)/(4n)$, $\omega_i = \omega_{i-2}$, $i = 2, \dots, n$, *is L-optimal for estimating the pair of coefficients* $\beta_{2\lfloor m/2\rfloor}$, $\beta_{4\lfloor m/2\rfloor}$. *Moreover,* $trLM^+(\xi^*_{(2\lfloor m/2\rfloor,\ 4\lfloor m/2\rfloor)}) = trLM^+(\xi^*_{(0,2\lfloor m/2\rfloor)}) = (\sqrt{5}/2) + (3/2)$.

(3) *The L-optimal design* $\xi^*_{(0,\ 2\lfloor m/2\rfloor)}$ *for estimating the pair of coefficients* β_0, $\beta_{2\lfloor m/2\rfloor}$ *coincides with the design* $\xi^*_{(2\lfloor m/2\rfloor,4\lfloor m/2\rfloor)}$ *defined in part* 2).

Proof. We only prove the first part of the theorem for the case where the degree m of the Fourier regression model is even. All other statements of the theorem are treated similarly. In this case $\lfloor m/2\rfloor = m/2$ and the design in part 1) of Theorem 3.1 can be rewritten as

$$\xi^*_{(m-1,\ 2m-1)} = \begin{pmatrix} -t_m & -t_{m-1} & \dots & -t_1 & t_1 & \dots & t_m \\ \frac{1}{2m} & \frac{1}{2m} & \dots & \frac{1}{2m} & \frac{1}{2m} & \dots & \frac{1}{2m} \end{pmatrix},$$

with $t_i = \lfloor i/2 \rfloor (2\pi/m) + (-1)^{(i-1)} x$, $x = (2/m)\arctan(\sqrt[4]{5})$. The proof is based on the characterization of L-optimal designs given in Theorem 2.1. Recall the definition of the matrix $M_s = M_s(\xi) \in \mathbb{R}^{m\times m}$ at (2.3), and let $m_{s[i,j]} = m_{s[i,j]}(\xi)$ denote the element of the matrix M_s in the position (i,j). We consider the system of equations

$$m_{s[j,\ m/2]} = 0, \quad j = 1,\ldots,m, \quad j \neq \frac{m}{2}, \tag{3.1}$$

$$m_{s[j,\ m]} = 0, \quad j = 1,\ldots,m-1, \tag{3.2}$$

$$m_{s[m/2,\ m/2]} = \sin^2\Big(\frac{m}{2}x\Big), \quad m_{s[m,m]} = \sin^2(mx). \tag{3.3}$$

We show below that the design $\xi^*_{(m-1,\ 2m-1)}$ satisfies (3.1)−(3.3). Under this assumption we obtain for the function $\varphi(t,\xi^*)$ in Theorem 2.1 the representation

$$\varphi(t,\xi^*) = f^T(t)M^+(\xi^*)LM^+(\xi^*)f(t) = \frac{\sin^2((m/2)t)}{\sin^4((m/2)x)} + \frac{\sin^2(mt)}{\sin^4(mx)}.$$

Now a straightforward calculation shows $trLM^+(\xi^*) = (\sqrt{5}/2) + (3/2)$, and it is easy to prove that (2.5) is satisfied calculating the solution of the equation$[(\partial\varphi(t,\xi^*))/(\partial t)] = 0$. Moreover, a similar calculation shows that $trLM^+(\xi^*) = \varphi(t_i,\xi^*)$ $= [(\sin^2(m/2)x)] + [1/(\sin^2(mx))] = (\sqrt{5}/2) + (3/2)$, and it remains to show (3.1)−(3.3). For a proof, we note that

$$m_{s[j,\ m/2]} = 2\sum_{k=1}^{m} \sin(jt_k)\sin(\frac{m}{2}t_k)\omega_k = \frac{1}{m}\sum_{k=1}^{m}\sin(jt_k)\sin(\frac{m}{2}t_k)$$

and that $\sin((m/2)t_k) = (-1)^{\lfloor (k-1)/2\rfloor}\sin((m/2)x)$, which yields $m_{s[j,\ m/2]} =$ $[\sin((m/2)x)/m]\sum_{k=1}^{m}(-1)^{\lfloor (k-1)/2\rfloor}\sin(jt_k)$. Next we prove that $\sum_{k=1}^{m}(-1)^{\lfloor (k-1)/2\rfloor}$ $\sin(jt_k) = 0$ if $j \neq (m/2)$. For this purpose we use the definition of t_k and obtain

$$\begin{aligned}
&\sum_{k=1}^{m}(-1)^{\lfloor (k-1)/2\rfloor}\sin(jt_k)\\
&\quad= \sin(jx) + \sin\Big(\frac{2\pi j}{m}\Big)\cos(jx) - \cos\Big(\frac{2\pi j}{m}\Big)\sin(jx)\\
&\qquad+(-1)\Big(\sin\Big(\frac{2\pi j}{m}\Big)\cos(jx) + \cos\Big(\frac{2\pi j}{m}\Big)\sin(jx)\Big) + \cdots\\
&\qquad+(-1)^{(m/2-1)}\Bigg(\sin\Big(\Big(\frac{m}{2}-1\Big)\frac{2\pi j}{m}\Big)\cos(jx) + \cos\Big(\Big(\frac{m}{2}-1\Big)\frac{2\pi j}{m}\Big)\sin(jx)\Bigg)\\
&\qquad+(-1)^{(j+1)+((m/2)-1)}\sin(jx)\\
&\quad= \sin(jx)\Bigg[2\sum_{k=0}^{(m/2)-1}(-1)^k\cos(\frac{2\pi kj}{m}) + (-1)^{j+(m/2)} - 1\Bigg].
\end{aligned}$$

Now, applying standard trigonometric formulae, it is easy to calculate the sum in the last expression if $j \neq (m/2)$,

$$
\begin{aligned}
2 &\sum_{k=0}^{(m/2)-1} (-1)^k \cos\Big(\frac{2\pi kj}{m}\Big) \\
&= \frac{1}{\cos(j\pi/m)}\Bigg(\Big(\cos\Big(\frac{j\pi}{m}\Big) + \cos\Big(\frac{3j\pi}{m}\Big)\Big) - \Big(\cos\Big(\frac{3j\pi}{m}\Big) + \cos\Big(\frac{5j\pi}{m}\Big)\Big) \\
&\quad + \cdots + (-1)^{(m/2)-1}\Big(\cos\Big(\frac{(m-3)j\pi}{m}\Big) + \cos\Big(\frac{(m-1)j\pi}{m}\Big)\Big)\Bigg) \\
&= 1 + (-1)^{j+(m/2)-1}.
\end{aligned}
$$

It follows that $\sum_{k=1}^{m} (-1)^{\lfloor (k-1)/2 \rfloor} \sin(jt_k) = \sin(jx)\,(1 + (-1)^{(m/2)-1+j} + (-1)^{j+(m/2)} - 1) = 0$, which shows that (3.1) is satisfied. The equality $m_{s[j,m]} = 0$ $(j \neq m)$ follows by similar arguments, that is $m_{s[j,m]} = \sin(jx)$ $(2\sum_{k=0}^{(m/2)-1} \cos(2\pi kj/m) + (-1)^j - 1) = 0$. Finally we obtain (3.3) by a direct calculation. This completes the proof of Theorem 3.1.

Remark 3.1. Note that Theorem 3.1 holds for trigonometric regression models of degree $m > 3$. If $m = 2$, the L-optimal design $\xi^*_{(1,3)}$ has equal masses at the points $-\pi+x$, $-x$, x, $\pi-x$, where $x = \arctan(\sqrt[4]{5})$ and $trLM_s^{-1}(\xi^*) = (3+\sqrt{5})/2$. For any $\alpha \in [0, (5-\sqrt{5})/8]$, the L-optimal design $\xi^*_{(0,2)}$ has masses $[(5-\sqrt{5})/8]-\alpha$, $(\sqrt{5}-1)/8$, $(5-\sqrt{5})/8$, $(\sqrt{5}-1)/8$, and α at the points $-\pi$, $-(\pi/2)$, 0, $\pi/2$, and π, respectively, with $trLM_c^{-1}(\xi^*) = (3+\sqrt{5})/2$. If $m = 3$, the L-optimal design $\xi^*_{(1,3)}$ has equal masses at the points $-\pi+x$, $-x$, x, $\pi-x$, where $x = \arctan(\sqrt[4]{5})$ and $trLM_s^{+}(\xi^*) = (3+\sqrt{5})/2$, while the L-optimal design $\xi^*_{(0,2)}$ has masses $(1/2)-2z-\alpha$, z, z, $(1/2)-2z$, z, , z, and α at the points $-\pi$, $-\pi+x$, $-x$, 0, x, $\pi-x$, and π, respectively, where $z \approx 0.15195067$, $x \approx 0.932928804$, and $trLM_c^{-1}(\xi^*) \approx 2.77004565$. The optimality of the designs can be easily checked with Theorem 2.1.

Remark 3.2. Note that, by Theorem 3.1, the sum of variances of the estimates for the corresponding coefficients in a Fourier regression model of degree $m > 3$ is $(\sqrt{5}/2) + (3/2)$ for the L-optimal design, while the D-optimal design yields 4 for this sum. Thus the reduction in the sum of variances obtained by the L-optimal design is approximately 35%.

There are two other cases where L-optimal designs for the trigonometric regression model (1.1) can be constructed explicitly, see the following two theorems. For the sake of brevity only a proof of Theorem 3.3 is given here.

Table 1. The solutions of x and z of the the system (3.4). The L-optimal design for estimating the specific pair is specified in the first part of Theorem 3.2.

$m = 3k$	$\{2k-1, 4k-1\}$	$\{2k-1, 6k-1\}$	$\{4k-1, 6k-1\}$
x	$\arctan(\sqrt[4]{5})$	0.6476	$3\pi/10$
z	$1/4$	0.14	$(3-\sqrt{5})/4$

Table 2. The solutions of x and z of the the system (3.5). The L-optimal design for estimating the specific pair is specified in the first part of Theorem 3.2.

$m = 3k$	$\{0, 2k\}$	$\{0, 4k\}$	$\{0, 6k\}$	$\{2k, 4k\}$	$\{2k, 6k\}$	$\{4k, 6k\}$
x	0.9329	$\pi/2$	$\pi/3$	1.1177	0.9232	1.1668
z	0.1519	$1/4$	$1/6$	0.1258	0.14	0.1478

Theorem 3.2. *Consider the trigonometric regression model* (1.1) *with* $m = 3k$. *The design*

$$\xi^* = \begin{pmatrix} -t_m & -t_{m-1} & \cdots & -t_1 & t_1 & \cdots & t_m \\ \omega_m & \omega_{m-1} & \cdots & \omega_1 & \omega_1 & \cdots & \omega_m \end{pmatrix}$$

with $t_1 = (\pi/2k) - (x/k)$, $t_2 = \pi/2k$, $t_3 = (\pi/2k) + (x/k)$, $t_i = t_{i-3} + (\pi/k)$, $i = 4, 5 \ldots, m$, $\omega_1 = z/k$, $\omega_2 = (1-4z)/2k$, $\omega_3 = z/k$, $\omega_i = \omega_{i-3}$, $i = 4, 5 \ldots, m$, *is L-optimal for estimating one of the pairs of the coefficients* $\{\beta_{2k-1}, \beta_{4k-1}\}$, $\{\beta_{2k-1}, \beta_{6k-1}\}$, $\{\beta_{4k-1}, \beta_{6k-1}\}$, *where only the values x and z depend on the particular pair under consideration, and satisfy*

$$\frac{\partial trLM_s^{-1}(\xi^*)}{\partial x} = 0 \qquad \frac{\partial trLM_s^{-1}(\xi^*)}{\partial z} = 0. \tag{3.4}$$

Similarly, for any $\alpha \in [0, \omega_m]$ *the design*

$$\xi^* = \begin{pmatrix} -\pi & -t_{m-1} & \cdots & -t_1 & 0 & t_1 & \cdots & t_{m-1} & \pi \\ \omega_m - \alpha & \omega_{m-1} & \cdots & \omega_1 & \omega_0 & \omega_1 & \cdots & \omega_{m-1} & \alpha \end{pmatrix}$$

with $t_0 = 0$, $t_1 = x/k$, $t_2 = (\pi - x)/k$, $t_i = t_{i-3} + (\pi/k)$, $i = 3, 4 \ldots, m-1$, $\omega_0 = (1-4z)/2k$, $\omega_1 = z/k$, $\omega_2 = z/k$, $\omega_i = \omega_{i-3}$, $i = 3, 4 \ldots, m$, *is L-optimal for estimating one of the pairs* $\{\beta_0, \beta_{2k}\}$, $\{\beta_0, \beta_{4k}\}$, $\{\beta_0, \beta_{6k}\}$, $\{\beta_{2k}, \beta_{4k}\}$, $\{\beta_{2k}, \beta_{6k}\}$, $\{\beta_{4k}, \beta_{6k}\}$, *where only the values x and z depend on the particular pair under consideration and satisfy*

$$\frac{\partial trLM_c^{-1}(\xi^*)}{\partial x} = 0 \qquad \frac{\partial trLM_c^{-1}(\xi^*)}{\partial z} = 0. \tag{3.5}$$

Some numerical values for the parameters x and z in Theorem 3.2 are presented in Tables 1 and 2.

Theorem 3.3. *Consider the trigonometric regression model* (1.1) *with* $m = 4k$. *The design*

$$\xi^* = \begin{pmatrix} -t_m & -t_{m-1} & \cdots & -t_1 & t_1 & \cdots & t_m \\ \omega_m & \omega_{m-1} & \cdots & \omega_1 & \omega_1 & \cdots & \omega_m \end{pmatrix}$$

with $t_1 = x_1/k$, $t_2 = x_2/k$, $t_3 = (\pi - x_2)/k$, $t_4 = (\pi - x_1)/k$, $t_i = t_{i-4} + \pi/k$, $i = 5, 6 \ldots, m$, $\omega_1 = z_1/k$, $\omega_2 = (1 - 4z_1)/m$, $\omega_3 = (1 - 4z_1)/m$, $\omega_4 = z_1/k$, $\omega_i = \omega_{i-4}$, $i = 5, 6 \ldots, m$, *is* L*-optimal for estimating one of the pairs of coefficients* $\{\beta_{2k-1}, \beta_{4k-1}\}$, $\{\beta_{2k-1}, \beta_{6k-1}\}$, $\{\beta_{2k-1}, \beta_{8k-1}\}$, $\{\beta_{4k-1}, \beta_{6k-1}\}$, $\{\beta_{4k-1}, \beta_{8k-1}\}$, $\{\beta_{6k-1}, \beta_{8k-1}\}$, *where only the values* x_1, x_2 *and* z_1 *depend on the particular pair under consideration and satisfy*

$$\frac{\partial trLM_s^{-1}(\xi^*)}{\partial x_1} = 0, \quad \frac{\partial trLM_s^{-1}(\xi^*)}{\partial x_2} = 0, \quad \frac{\partial trLM_s^{-1}(\xi^*)}{\partial z_1} = 0. \tag{3.6}$$

Similarly, if $n = 5m/4$, *then for any* $\alpha \in [0, \omega_n]$ *the design*

$$\xi^* = \begin{pmatrix} -\pi & -t_{n-1} & \cdots & -t_1 & 0 & t_1 & \cdots & t_{n-1} & \pi \\ \omega_n - \alpha & \omega_{n-1} & \cdots & \omega_1 & \omega_0 & \omega_1 & \cdots & \omega_{n-1} & \alpha \end{pmatrix}$$

with $t_0 = 0$, $t_1 = x_1/k$, $t_2 = x_2/k$, $t_3 = (\pi - x_2)/k$, $t_4 = (\pi - x_1)/k$, $t_i = t_{i-5} + \pi/k$, $i = 5, 6 \ldots, n-1$, $\omega_0 = (1 - 4z_1 - 4z_2)/2k$, $\omega_1 = \omega_4 = z_1/k$, $\omega_2 = \omega_3 = z_2/k$, $\omega_i = \omega_{i-5}$, $i = 5, 6, \ldots, n$, *is optimal for estimating any of the pairs of the coefficients* $\{\beta_0, \beta_{4k}\}$, $\{\beta_0, \beta_{6k}\}$, $\{\beta_0, \beta_{8k}\}$, $\{\beta_{2k}, \beta_{4k}\}$, $\{\beta_{2k}, \beta_{6k}\}$, $\{\beta_{2k}, \beta_{8k}\}$, $\{\beta_{4k}, \beta_{6k}\}$, $\{\beta_{4k}, \beta_{8k}\}$, $\{\beta_{6k}, \beta_{8k}\}$, *where only the values* x_1, x_2, z_1, *and* z_2 *depend on the particular pair under consideration and satisfy*

$$\begin{aligned} \frac{\partial trLM_c^{-1}(\xi^*)}{\partial x_1} = 0, \quad \frac{\partial trLM_c^{-1}(\xi^*)}{\partial x_2} = 0 \\ \frac{\partial trLM_c^{-1}(\xi^*)}{\partial z_1} = 0, \quad \frac{\partial trLM_c^{-1}(\xi^*)}{\partial z_2} = 0. \end{aligned} \tag{3.7}$$

The numerical values of the quantities x_1, x_2, z_1, and z_2 are listed in the Tables 3 and 4.

Proof of Theorem 3.3. We only prove the first part of Theorem 3.3, the second part is treated similarly. We begin with the case $k = 1$, i.e., $m = 4$, for which it is easy to check by direct calculations that the design ξ^* defined in the first part of Theorem 3.3 is L-optimal. The corresponding numerical values of the quantities x_1, x_2, and z_1 can be found as the solution of (3.6). Now let $k \geq 2$ and $m = 4k$.

Table 3. The solutions x_1, x_2 and z_1 of the the system (3.6). The L-optimal design for estimating the specific pair of coefficients in the Fourier regression model (1.1) is specified in the first part of Theorem 3.3.

$m=4k$	$\{2k-1,4k-1\}$	$\{2k-1,6k-1\}$	$\{2k-1,8k-1\}$
x_1	$\pi/4$	0.6476	0.4845
x_2	$\pi/2$	$\pi/2$	1.1912
z_1	$(6-\sqrt{6})/20$	0.14	0.0909
$m=4k$	$\{4k-1,6k-1\}$	$\{4k-1,8k-1\}$	$\{6k-1,8k-1\}$
x_1	0.7338	$\arctan(\sqrt[4]{5})/2$	0.4523
x_2	1.3884	$(\pi-\arctan(\sqrt[4]{5}))/2$	1.2566
z_1	0.168	1/8	0.1417

Table 4. The solutions of x_1, x_2, z_1 and z_2 of the the system (3.7). The L-optimal design for estimating the specific pair of coefficients in the Fourier regression model (1.1) is specified in the second part of Theorem 3.3.

$m=4k$	{0,4k}	{0,6k}	{0,8k}	{2k,4k}	{2k,6k}
x_1	$\pi/4$	$\pi/3$	$\pi/4$	$\pi/4$	0.9232
x_2	$\pi/2$	$\pi/3$	$\pi/2$	$\pi/2$	0.9232
z_1	0.0863	1/6	1/8	$(2\sqrt{2}-1)/28$	0.07
z_2	0.0773	1/6	1/8	$(4-\sqrt{2})/28$	0.07
$m=4k$	{2k,8k}	{4k,6k}	{4k,8k}	{6k,8k}	
x_1	0.7132	1.0157	$\pi/4$	0.8814	
x_2	$\pi/2$	$\pi/2$	$\pi/2$	$\pi/2$	
z_1	0.033	0.0708	0.0863	0.0477	

We consider the system of equations

$$\begin{cases} m_{s[k,j]} &= 0, \quad j=1,\ldots,4k, \quad j\neq k,\ j\neq 3k, \\ m_{s[4k,j]} &= 0, \quad j=1,\ldots,4k-1, \quad j\neq 2k, \\ m_{s[k,k]} &= 4\Big(z_1\sin^2(x_1)+\Big(\frac{1}{4}-z_1\Big)\sin^2(x_2)\Big), \\ m_{s[k,3k]} &= 4\Big(z_1\sin(x_1)\sin(3x_1)+\Big(\frac{1}{4}-z_1\Big)\sin(x_2)\sin(3x_2)\Big), \\ m_{s[4k,2k]} &= 4\Big(z_1\sin(2x_1)\sin(4x_1)+\Big(\frac{1}{4}-z_1\Big)\sin(2x_2)\sin(4x_2)\Big), \\ m_{s[4k,4k]} &= 4\Big(z_1\sin^2(4x_1)+\Big(\frac{1}{4}-z_1\Big)\sin^2(4x_2)\Big), \end{cases} \tag{3.8}$$

where $m_{s[i,j]}=m_{s[i,j]}(\xi^*)$ is the element of the matrix $M_s(\xi^*)\in R^{m\times m}$ in the ith row and jth column. We prove below that these equalities are satisfied for the design ξ^*. In this case it follows that the quantities $trLM_s^{-1}(\xi^*)$ and $\varphi(t,\xi^*) = f^T(t)M^+(\xi^*)LM^+(\xi^*)f(t)$ in Theorem 2.1 are independent of the value k (note that the matrix L is a given diagonal matrix where the non-vanishing entries depend on the particular pair of parameters under consideration). Consequently it is sufficient to prove Theorem 3.3 in the case $k=1$, which has been done in the previous paragraph.

In order to prove that the equalities (3.8) are satisfied we note that for $i=1,\ldots,4k-1$, $i \neq k,2k,3k$, we have

$$\begin{aligned} m_{s[k,i]} &= 2\sum_{j=1}^{4k} \sin(kt_j)\sin(it_j)\omega_j \\ &= 2\sin(x_1)\frac{z_1}{k}\sum_{j=1}^{k}(-1)^{j-1}\Big(\sin(it_{4j-3})+\sin(it_{4j})\Big) \\ &\quad +2\sin(x_2)\Big(\frac{1}{4}-z_1\Big)\frac{1}{k}\sum_{j=1}^{k}(-1)^{j-1}\Big(\sin(it_{4j-2})+\sin(it_{4j-1})\Big). \end{aligned}$$

For the first sum on the right side we obtain $\sum_{j=1}^{k}(-1)^{j-1}(\sin(it_{4j-3})+\sin(it_{4j})) = 0$. Similarly, it follows (substituting x_1 for x_2) that the second sum also vanishes, which implies $m_{s[k,i]}=0$ for $i=1,\ldots,4k-1$, $i\neq k,2k,3k$. For the element $m_{s[4k,i]}$ we find for $i=1,\ldots,4k-1$, $i\neq k,2k,3k$,

$$\begin{aligned} m_{s[4k,i]} &= 2\sin(x_1)\frac{z_1}{k}\sum_{j=1}^{k}\Big(\sin(it_{4j-3})-\sin(it_{4j})\Big) \\ &\quad +2\sin(x_2)\Big(\frac{1}{4}-z_1\Big)\frac{1}{k}\sum_{j=1}^{k}\Big(\sin(it_{4j-2})-\sin(it_{4j-1})\Big). \end{aligned}$$

A straightforward calculation now yields

$$\begin{aligned} &\sum_{j=1}^{k}\Big(\sin(it_{4j-3})-\sin(it_{4j})\Big) \\ &= \frac{\Big(\cos\big(\frac{ix_1-i\pi}{k}-i\pi\big)-\cos\big(\frac{ix_1-i\pi}{k}+i\pi\big)+\cos\big(\frac{ix_1}{k}-i\pi\big)-\cos\big(\frac{ix_1}{k}+i\pi\big)\Big)}{2\sin\big(\frac{i\pi}{k}\big)} = 0, \end{aligned}$$

and the same arguments show that the second sum also vanishes, which implies $m_{s[4k,i]}=0$. To conclude the proof it remains to calculate $m_{s[ik,jk]}$, $i,j=1,\ldots,4$, for which we obtain

$$
\begin{aligned}
m_{s[ik,jk]} = {} & \frac{2z_1}{k}\sum_{r=1}^{k}\Big(\sin(i(x_1+(r-1)\pi))\sin(j(x_1+(r-1)\pi)) \\
& +\sin(i(-x_1+r\pi))\sin(j(-x_1+r\pi))\Big) \\
& +\Big(\frac{1}{4}-z_1\Big)\frac{2\sin(ix_2)\sin(jx_2)}{k}\sum_{r=1}^{k}\Big((-1)^{(i+j)(r-1)}+(-1)^{(i+j)r}\Big) \\
= {} & \Big(\frac{2\sin(ix_1)\sin(jx_1)z_1}{k}+\Big(\frac{1}{4}-z_1\Big)\frac{2\sin(ix_2)\sin(jx_2)}{k}\Big) \\
& \times\sum_{r=1}^{k}\Big((-1)^{(i+j)(r-1)}+(-1)^{(i+j)r}\Big).
\end{aligned}
$$

From this representation it is obvious that $m_{s[ik,jk]}$, $i,j = 1,\dots,4$, have the values specified by the system of equations, and the theorem has been proved.

Remark 3.3. Note that for any k with $m/2 < k \le m$, and for any $\beta \in [0,(1/2k)]$, the design $\xi^*_{(0,2k)}$ with equal masses at the points $-\pi, -\pi+(\pi/k),\dots, pi+((2k-1)/k)\pi, \pi$ is L-optimal for estimating the pair of coefficients $\{\beta_0,\beta_{2k}\}$. Moreover, in this case it follows that $trLM_c^{-1}(\xi^*_{(0,2k)}) = 2$ (see Dette and Melas (2003, Lemma 2.3)).

4. Examples

In this section we present several examples that illustrate the theoretical results obtained in Section 3.

Example 4.1. We present the L-optimal design for estimating the coefficients β_3, β_7 (i.e., the coefficients of $\sin(2t)$ and $\sin(4t)$) in the trigonometric regression model of degree 4. The L-optimal design $\xi^*_{(3,7)}$ is directly obtained from Theorem 3.1 and has equal masses at the points $-\pi+x$, $-(\pi/2)-x$, $-(\pi/2)+x$, $-x$, x, $(\pi/2)-x$, $(\pi/2)+x$, $\pi-x$, where $x = 1/2\arctan(\sqrt[4]{5}) \approx 0.49068$. The corresponding support points of the design $\xi^*_{(3,7)}$ are depicted in the left part of Figure 1.

A straightforward calculation shows that the function $\varphi(t,\xi^*_{(3,7)})$ is given explicitly by

$$
\begin{aligned}
\varphi(t,\xi^*_{(3,7)}) &= f^T(t)M^+(\xi^*_{(3,7)})LM^+(\xi^*_{(3,7)})f(t) \\
&= \frac{(1+\sqrt{5})^2}{5}\sin^2(2t)+\frac{(3+\sqrt{5})^2}{20}\sin^2(4t).
\end{aligned}
$$

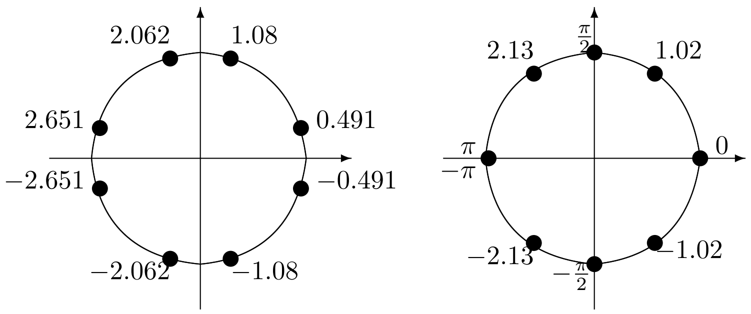

Figure 1. Left part: Support points of the L-optimal design for estimating the coefficients β_3, β_7 in the Fourier regression model of degree 4. Right part: support points of the L-optimal design for estimating β_4, β_6.

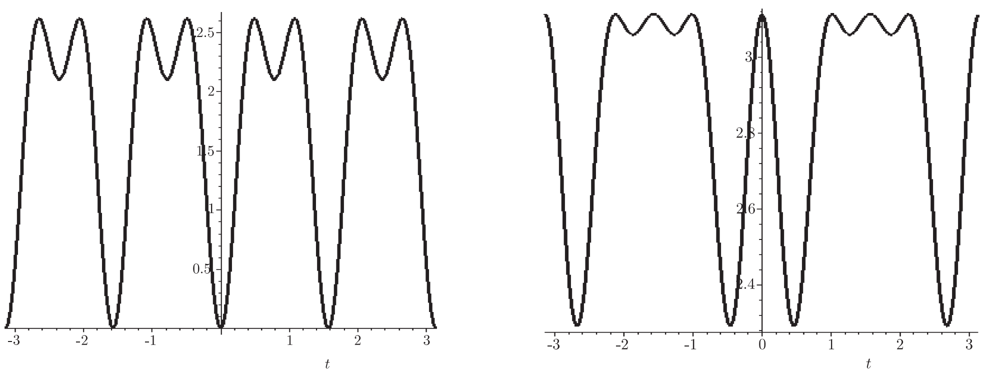

Figure 2. Left part: the function $\varphi(t, \xi^*_{(3,7)})$ defined in Example 4.1. Right part: The function $\varphi(t, \xi^*_{(4,6)})$ defined in Example 4.2.

Then (2.6) is checked by a direct calculation, and the function $\varphi(t, \xi^*_{(3,7)})$ is depicted in the right part of Figure 2.

Example 4.2. Consider the trigonometric regression model of degree $m = 4$, and use Theorem 3.3 to determine the L-optimal design for estimating the pair of coefficients β_4 and β_6 that correspond to the terms $\cos(2t)$ and $\cos(3t)$. The L-optimal design $\xi^*_{(4,6)}$ for estimating these coefficients has masses $0.175 - \alpha$, 0.09, 0.145, 0.09, 0.175, 0.09, 0.145, 0.09, α at the points $-\pi$, -2.13, $-(\pi/2)$, -1.02, 0, 1.02, $\pi/2$, 2.13, π where $\alpha \in [0, 0.175]$. The support points of the optimal design are depicted in the right part of Figure 1. We finally note that a straightforward calculation yields, for the function $\varphi(t, \xi^*_{(4,6)})$ in Theorem 2.1, $\varphi(t, \xi^*_{(4,6)}) = f^T(t)M^+(\xi^*_{(4,6)})LM^+(\xi^*_{(4,6)})f(t) = 2.851 - 0.262\cos(2t) + 0.116\cos(4t) + 0.262\cos(6t) + 0.147\cos(8t)$.

Acknowledgements

The support of the Deutsche Forschungsgemeinschaft (SFB 475, "Komplexitätsreduktion in multivariaten Datenstrukturen") is gratefully acknowledged. The work of the authors was also supported in part by an NIH grant award IR01GM072876:01A1 and by the BMBF project SKAVOE.

References

Biedermann, S., Dette, H. and Hoffmann, P. (2007). Constrained optimal discriminating designs for Fourier regression models. *Ann. Statist. Math.*, in press.

Currie, A. J., Ganeshanandam, S., Noition, D. A., Garrick, D., Shelbourne, C. J. A. and Oragguzie, N. (2000). Quantiative evalution of apple (Malus x domestica Borkh.) fruit shape by principle component analysis of Fourier descriptors. *Euphytica* **111**, 219-227.

Dette, H. and Haller, G. (1998). Optimal designs for the identification of the order of a Fourier regression. *Ann. Statist.* **26**, 1496-1521.

Dette, H. and Melas, V. B. (2002). E-optimal designs for Fourier regression models. *Math. Methods Statist.* **11**, 259-296.

Dette, H. and Melas, V. B. (2003). Optimal designs for estimating individual coefficients in Fourier regression models. *Ann. Statist.* **31**, 1669-1692.

Dette, H., Melas, V. B. and Pepelysheff, A. (2002). D-Optimal designs for trigonometric regression models on a partial circle. *Ann. Inst. Statist. Math.* **54**, 945-959.

Dette, H., Melas, V. B. and Shpilev, P. V. (2007). Optimal designs for estimating the coefficients of the lower frequencies in trigonometric regression models. *Ann. Inst. Statist. Math.* **59**, 655-673.

Ermakov, M. and Zhigljavsky, A. (1987). *Mathematical Theory of Optimal Experiment.* Nauka, Moscow. (in Russian).

Karlin, S. and Studden, W. J. (1966). *Tchebycheff Systems: With Applications in Analysis and Statistics.* Interscience, New York.

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

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

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

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

McCool, J. I. (1979). Systematic and random errors in least squares estimation for circular contours. *Precision Engineering* **1**, 215-220.

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

Rao, C. R. (1968). *Linear Statistical Inference and Its Applications.* Wiley, New York.

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

Young, J. C. and Ehrlich, R. (1977). Fourier biometrics: harmonic amplitudes as multivariate shape descriptors. *Systematic Zoology* **26**, 336-342.

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

Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany.

E-mail: holger.dette@rub.de

St. Petersburg State University, Department of Mathematics, St. Petersburg, Russia.

E-mail: v.melas@pobox.spbu.ru

St. Petersburg State University, Department of Mathematics, St. Petersburg, Russia.

E-mail: pitshp@hotmail.com

(Received January 2008; accepted May 2008)