Journal of Statistical Planning and Inference 143 (2013) 1186–1194

Contents lists available at SciVerse ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

D-optimal designs for combined polynomial and trigonometric regression on a partial circle

Fu-Chuen Chang *, Chin-Han Li

Department of Applied Mathematics, National Sun Yat-sen University, Kaohsiung 804, Taiwan, ROC

ARTICLE INFO

Article history:
Received 22 October 2008
Received in revised form
8 June 2012
Accepted 31 January 2013
Available online 16 February 2013

Keywords:
D-optimal design
Implicit function theorem
Polynomial regression
Recursive algorithm
Taylor expansion
Trigonometric regression

ABSTRACT

Consider the *D*-optimal designs for a combined polynomial and trigonometric regression on a partial circle. It is shown that the optimal design is equally supported and the structure of the optimal design depends only on the length of the design interval and the support points are analytic functions of this parameter. Moreover, the Taylor expansion of the optimal support points can be determined efficiently by a recursive procedure. Examples are presented to illustrate the procedures for computing the optimal designs.

1. Introduction

Consider the following combined polynomial of degree d and trigonometric of order m regression model ($PT(d,m)$) (see Graybill, 1976, p. 324):

$$y(x) = \sum_{i=0}^{d} \beta_i x^i + \sum_{i=1}^{m} [\alpha_{2i-1} \sin(ix) + \alpha_{2i} \cos(ix)] + \varepsilon(x), \tag{1.1}$$

where $x \in I = [b,c]$ with $-\infty < b$, $c < \infty$ and $\varepsilon(x)$ is a random error component. The errors are assumed to have mean zero and unknown variance σ^2. Additionally we assume that the errors are uncorrelated. The response $y(x)$ is a random variable with mean $\sum_{i=0}^{d} \beta_i x^i + \sum_{i=1}^{m} [\alpha_{2i-1} \sin(ix) + \alpha_{2i} \cos(ix)]$ and variance σ^2.

This model is widely used in situations where the scatter diagram of y versus x indicates periodicity or cyclic behavior in the data, adding trigonometric terms to the polynomial model may be very beneficial. This benefit has been noted by both Graybill (1976) and Eubank and Speckman (1990). It is a natural extension of the polynomial regression models ($PT(d,0)$) which is widely used in situations where the response is curvilinear, and trigonometric regression models ($PT(0,m)$) which are widely used to describe periodic phenomena.

The problem of determining optimal designs for polynomial regression models has been investigated by several authors (see Hoel, 1958; Karlin and Studden, 1966a; Chang and Lin, 1997; Antille et al., 2003; Chang, 2005, among many others). For related works on optimal designs for trigonometric regression models see Hoel (1965), Lau and Studden (1985),

* Corresponding author. Tel.: +886 7 5252000x3823; fax: +886 7 5253809.
E-mail addresses: changfc@math.nsysu.edu.tw, fuchuen@gmail.com (F.-C. Chang).

0378-3758/$ - see front matter
http://dx.doi.org/10.1016/j.jspi.2013.01.013

Pukelsheim (1993, p. 241), Riccomagno et al. (1997), Dette and Haller (1998), Wu (2002), Dette et al. (2004), and Melas (2006) among many others.

Most authors focus on the design interval $[-\pi,\pi]$, (see e.g. Hoel, 1965; Karlin and Studden, 1966b, p. 347; Fedorov, 1972, p. 94; Dette and Haller, 1998), but Hill (1978) and Kitsos et al. (1988) show that in many applications it is impossible to take observations on the full circle $[-\pi,\pi]$. We refer to Kitsos et al. (1988) for a real example, who studied a design problem in rhythmometry involving circadian rhythm exhibited by peak expiratory flow, for which the design interval has to be restricted to a partial cycle of the complete 24-h period.

Apparently, there is no literature on how to determine the optimal designs for $PT(d,m)$ on a partial circle with $d \ge 1$ and $m \ge 1$. A special case of combined linear $(d=1)$ and trigonometric model of order m on the full circle $[-\pi,\pi]$ was considered in Wu (2003). He proved that the equispaced minimum-support designs are D-optimal. A design ξ is called minimum-support for $PT(d,m)$ if its number of support points equals the number of unknown parameters $d+2m+1$.

The purpose of this paper is to determine the D-optimal designs for combined polynomial and trigonometric regression on a partial circle, $c-b \le 2\pi$. The optimal designs are characterized in a functional way. It can be easily generalized to the case $c-b > 2\pi$. The approach is based on the fact that the optimal support points are analytic functions of the length of the design interval. Moreover, the Taylor expansions of the optimal support points can be computed recursively. It was used in Dette et al. (2002) and Chang (2005) to derive the D-optimal designs for $PT(0,m)$ and weighted $PT(d,0)$ models, respectively.

This paper is organized in the following way. In Section 2, the structure of the D-optimal designs as a approaches to 0 is given. In Section 3, we show that the D-optimal support points are analytic functions of a. A recursive formula is given for computing Taylor polynomials of D-optimal support points. Finally, in Section 4 two examples are used to illustrate the Taylor expansion method established in Section 3. Concluding comments are given in Section 5. All the proofs of lemmas and theorems are deferred to an Appendix.

2. Preliminary results

Consider the combined polynomial and trigonometric regression model (1.1), define $\beta=(\beta_0,\ldots,\beta_d,\alpha_1,\ldots,\alpha_{2m})^T$ as the vector of unknown parameters of interest and

$$f(x)=(1,x,\ldots,x^d,\sin x,\cos x,\ldots,\sin(mx),\cos(mx))^T \tag{2.1}$$

as the vector of regression functions. An approximate design ξ is a probability measure with finite support on the design interval I. For the uncorrelated observations the information matrix of a design ξ for the parameter β is defined by

$$M(\xi)=\int_I f(x)f^T(x)\,d\xi(x).$$

A design ξ^* is called approximate D-optimal for β if ξ^* maximizes the determinant of the information matrix $M(\xi)$ among the set of all approximate designs on I. For more details on the theory of optimal designs see Fedorov (1972), Silvey (1980), and Pukelsheim (1993). Furthermore, the D-optimality of a design can be checked by D-Equivalence Theorem (Kiefer and Wolfowitz, 1960). The theorem states that a design ξ^* is D-optimal for β if and only if

$$d(x,\xi^*)=f^T(x)M^{-1}(\xi^*)f(x) \tag{2.2}$$

$$d(x,\xi^*) \le d+2m+1 \tag{2.3}$$

for all $x \in I$. Here the equality holds if x belongs to the support of ξ^*.

It is well known that the determinant of information matrix of a design for both polynomial $PT(d,0)$ and trigonometric $PT(0,m)$ models are invariant under translation of support points (see Fedorov, 1972, Theorem 2.2.4; Dette et al., 2002, Lemma 2.1). First we show that it also holds true for $PT(d,m)$ models, $d \ge 1$ and $m \ge 1$. Let

$$\psi=\begin{Bmatrix} x_1 & \cdots & x_n \\ w_1 & \ldots & w_n \end{Bmatrix} \tag{2.4}$$

denote a design on $[b,c]$ with $b \le x_1 < \cdots < x_n \le c$ and $\sum_{i=1}^n w_i=1$, $0 < w_1,\ldots,w_n < 1$. Consider an affine transformation: $z=x-(b+c)/2$ which maps the interval $[b,c]$ onto the symmetric interval $[-a,a]$ and define

$$\xi=\begin{Bmatrix} z_1 & \cdots & z_n \\ w_1 & \ldots & w_n \end{Bmatrix}, \tag{2.5}$$

where $a=(c-b)/2$ and $z_i=x_i-(b+c)/2$, $i=1,\ldots,n$.

Lemma 2.1. *Let $M(\psi)$ and $M(\xi)$ denote the information matrices of the designs ψ and ξ defined by (2.4) and (2.5) for $PT(d,m)$, respectively, then*

$$\det M(\xi)=\det M(\psi). \tag{2.6}$$