Statistica Sinica **8**(1998), 1271-1274

D-OPTIMAL DESIGNS FOR POLYNOMIAL REGRESSION WITH WEIGHT FUNCTION $x/(1+x)$

L. Imhof, O. Krafft and M. Schaefer

Technical University of Aachen

Abstract: For polynomial regression with weight function $x/(1+x)$ for $x \in [0, a]$, a polynomial is presented whose zeros are the support points of the D-optimal approximate design.

Key words and phrases: D-optimal design, Heun-type differential equation, oscillatory matrices, weighted polynomial regression.

1. Introduction

If in the ordinary polynomial regression set-up a constant weight function or a Jacobi weight function is used, then the Stieltjes-Schur approach to maximizing a discriminant via an appropriate differential equation leads directly to a solution of the D-optimal design problem (e.g. Szegö (1975), p. 140). For other weight functions the design problem appears to be less tractable. For instance, Huang, Chang and Wong (1995) and Chang and Lin (1997) investigated some non-classical weight functions and, making use of the Stieltjes-Schur approach, arrived at appropriate differential equations and a corresponding eigenvalue problem. In this note we consider the weight function $x/(1+x)$, and solve the associated eigenvalue problem explicitly. A polynomial whose zeros describe the optimal design is specified. For varying degree d these polynomials do not form an orthogonal set. Properties of oscillatory matrices are the main tool to identify the solution.

It should be mentioned that we came to this special weight function from working on the still unsolved problem of characterizing A-optimal designs. In this context it can easily be shown that the support of an A-optimal design for the vector $(1, x, \ldots, x^{2d+1})^T$ of regression functions on $[-1, 1]$ can simply be calculated from that of an A-optimal design for $w^{1/2}(x)(1, x, \ldots, x^d)^T$ on [0,1] where $w(x) = x/(1+x)$.

2. Problem and Result

Consider the dth degree polynomial regression model where one has random variables $Y(x), x \in [0, a]$, with common variance σ^2 not depending on x and

$E(Y(x)) = \beta^T \mathbf{f}(x)$. Here $\mathbf{f}(x) = w^{1/2}(x)(1, x \ldots, x^d)^T$ with $w(x) = x/(1+x)$, and β is a $(d+1)$-dimensional vector of unknown parameters. An approximate design ξ, i.e. a probability measure on [0,a], is said to be D-optimal for this model (based on uncorrelated observations of $Y(x)$) if it maximizes det $\int \mathbf{f}(x)\mathbf{f}^T(x)d\xi(x)$.

Theorem. *The D-optimal design measure on* $[0, a]$ *puts probability* $(d+1)^{-1}$ *at the point a and at each of the d zeros of the polynomial* $g_d(x) = c_0 + \cdots + c_d x^d$, *where*

$$c_i = (-a)^{d-i} \binom{d+i}{i} \binom{d}{i} [2i(a+1) + 1 + \rho] \tag{1}$$

and $\rho = (4d(d+1)(a+1)+1)^{1/2}$.

Proof. Let ξ^* be any D-optimal design measure, existence being ensured by Lemma X.2.1 of Karlin and Studden (1966). It is readily verified that $\{1, w(x), xw(x), \ldots, x^{2d}w(x)\}$ is a Chebyshev system on $[0, a]$. By Theorem X.3.6 of Karlin and Studden (1966), ξ^* must have exactly $d+1$ support points, say $x_1^* < \cdots < x_{d+1}^*$. Therefore, ξ^* puts probability $(d+1)^{-1}$ at each of them (e.g. Lemma 5.1.3 in Silvey (1980)). Now if ξ is any design measure which puts probability $(d+1)^{-1}$ at $d+1$ points $x_1, \ldots, x_{d+1} \in [0, a]$, then det $\int \mathbf{f}(x)\mathbf{f}^T(x)d\xi(x)$ is proportional to

$$D(x_1, \ldots, x_{d+1}) := \Big(\prod_{k=1}^{d+1} w(x_k) \Big) \Big(\prod_{i<j} (x_i - x_j)^2 \Big).$$

Thus D is maximal at $(x_1^*, \ldots, x_{d+1}^*)$, and this implies that $x_1^* > 0$ and $x_{d+1}^* = a$. Moreover, the partial derivatives $\partial \log D/(\partial x_i), i = 1, \ldots, d$, must vanish at $(x_1^*, \ldots, x_{d+1}^*)$; that is,

$$\frac{1}{x_i^*(1+x_i^*)} + 2 \sum_{\substack{j=1 \\ j \neq i}}^{d+1} \frac{1}{x_i^* - x_j^*} = 0, \quad i = 1, \ldots, d.$$

Setting $F(x) = (x - x_1^*) \ldots (x - x_{d+1}^*)$, one has $F''(x_i^*)/F'(x_i^*) = 2\sum_{j \neq i}(x_i^* - x_j^*)^{-1}$. It follows that $H(x) := x(1+x)F''(x) + F'(x)$ vanishes at $x = x_1^*, \ldots, x_d^*$. Being a polynomial of degree $d+1, H$ must have yet another zero, say y^*, and so $H(x) = c(x - y^*)G(x)$, where $G(x) = (x - x_1^*) \ldots (x - x_d^*)$ and $c = d^2 + d$. Hence

$$x(1+x)F''(x) + F'(x) - c(x - y^*)G(x) = 0.$$

Now $F(x) = (x - a)G(x)$, so that $F'(x) = G(x) + (x-a)G'(x), F''(x) = 2G'(x) + (x-a)G''(x)$. It follows that G must satisfy the Heun-type differential equation

$$x(1+x)(x-a)G''(x) + (2x^2 + 3x - a)G'(x) - cxG(x) = -(1 + cy^*)G(x). \tag{2}$$

Writing $G(x) = \sum_{i=0}^{d} a_i x^i$ and equating the coefficients of equal powers of x on both sides of (2), one observes that the coefficient vector $\mathbf{a} = (a_0, \ldots, a_d)^T$ is a characteristic vector associated with the characteristic value $-(1 + cy^*)$ of the Jacobi matrix

$$\mathbf{A} = \left\| \begin{matrix} \beta_0 & \gamma_0 & 0 & \cdots & 0 & 0 \\ \alpha_1 & \beta_1 & \gamma_1 & \cdots & 0 & 0 \\ 0 & \alpha_2 & \beta_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \alpha_d & \beta_d \end{matrix} \right\|, \quad \text{where} \quad \begin{aligned} \alpha_i &= -(d+i)(d+1-i), \\ \beta_i &= i(2+a+i-ai), \\ \gamma_i &= -a(i+1)^2. \end{aligned}$$

A straightforward if somewhat lengthy calculation shows that the coefficient vector $\mathbf{c} = (c_0, \ldots, c_d)^T$ from (1) is a characteristic vector of $\mathbf{A}$ associated with the characteristic root $c + (\rho - 1)/2$.

Now consider the Jacobi matrix $\mathbf{B} := \delta\mathbf{I} - \mathbf{A}$, where $\mathbf{I}$ is the unit matrix of order $d{+}1$ and $\delta \in I\!R$ is chosen so large that every principal minor of $\mathbf{B}$ is positive. The sub- and super-diagonal entries of $\mathbf{B}$ are obviously positive. It follows that $\mathbf{B}$ is an oscillatory matrix, see Gantmacher (1959), pp. 103f. As $\mathbf{a}$ and $\mathbf{c}$ are characteristic vectors of $\mathbf{A}$, they are also characteristic vectors of $\mathbf{B}$. Besides, as all the roots of G are positive, the sequence $a_0, \ldots, a_d$ has, according to Descartes' rule of signs, exactly d sign changes, and the sequence $c_0, \ldots, c_d$ evidently has d sign changes too. It now follows from Theorem XIII.13 of Gantmacher (1959) that $\mathbf{a}$ and $\mathbf{c}$ are proportional, so that G and g_d have the same zeros. That is, the zeros of g_d and the point a are the support points of the unique D-optimal design measure.

Remarks. (a) From the recurrence formula for orthogonal polynomials (cf. Szegö (1975), Theorem 3.2.1) we see that the polynomials $g_d, d \in I\!N$, are not orthogonal.
(b) In a private communication Holger Dette provided a representation of g_d in terms of the Jacobi polynomials $P_d^{(\alpha,\beta)}(x)$ on [-1,1]. Setting

$$b_d = d[2d(a+1) + 1 - \rho]/[2(a+1)d],$$

it holds that

$$\begin{aligned} b_d g_d(x) = {} & [2d(a+1)+1](d+1)xP_{d-1}^{(1,1)}(2xa^{-1} - 1) \\ & - d(2d+1)P_d^{(0,0)}(2xa^{-1} - 1) - d\rho P_{d-1}^{(1,0)}(2xa^{-1} - 1). \end{aligned}$$

(c) Following a suggestion of a referee, we give a table of the optimal support points for the interval $[0, a] = [0, 2]$ and degrees $d \in \{1, \ldots, 7\}$. For comparison the corresponding points for the model with weight function $w(x) \equiv 1$ are given in brackets.

Table 1. Support points for $[0, a] = [0, 2], 1 \leq d \leq 7$. In brackets the Legendre-points.

d								
1	0.5	2						
	[0	2]						
2	0.2469	1.1961	2					
	[0	1.0	2]					
3	0.1479	0.7429	1.5293	2				
	[0	0.5528	1.4472	2]				
4	0.0985	0.5017	1.1118	1.6957	2			
	[0	0.3453	1.0	1.6547	2]			
5	0.0703	0.3608	0.8274	1.3533	1.7883	2		
	[0	0.2349	0.7148	1.2852	1.7651	2]		
6	0.0527	0.2717	0.6348	1.0783	1.5126	1.8446	2	
	[0	0.1698	0.5312	1.0	1.4688	1.8302	2]	
7	0.0409	0.2119	0.5008	0.8701	1.2646	1.6213	1.8812	2
	[0	0.1283	0.4083	0.7907	1.2093	1.5917	1.8717	2]

Acknowledgement

We thank a referee for several suggestions improving the presentation of the manuscript.

References

Chang, F.-C. and Lin, G.-C. (1997). *D*-optimal designs for weighted polynomial regression. *J. Statist. Plann. Inference* **62**, 317-331.

Gantmacher, F. R. (1959). *The Theory of Matrices.* Vol 2. Chelsea, New York.

Huang, M.-N. L., Chang, F.-C. and Wong, W. K. (1995). *D*-optimal designs for polynomial regression without an intercept. *Statist. Sinica* **5**, 441-458.

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

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

Szegö, G. (1975). *Orthogonal Polynomials.* Amer. Math. Soc., Providence, Rhode Island.

Institut für Statistik und Wirtschaftsmathematik, Rheinisch-Westfälische Technische Hochschule, Wüllnerstraβe 3, D-52056 Aachen, Germany.

(Received December 1997; accepted March 1998)