A note on optimal designs in weighted polynomial regression for the classical efficiency functions

Gérard Antille
Department of Econometrics
University of Geneva
UniMail, Bd du Pont-d'Arve 40
CH-1211 Geneva 4
Switzerland
email: Gerard.Antille@metri.unige.ch

Holger Dette
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany
email: holger.dette@ruhr-uni-bochum.de
FAX: +49 2 34 32 14 559

Anna Weinberg
Department of Econometrics
University of Geneva
UniMail, Bd du Pont-d'Arve 40
CH-1211 Geneva 4
Switzerland
email: Anna.Weinberg@metri.unige.ch

February 12, 2001

Abstract

In this note we consider the D-optimal design problem for the heteroscedastic polynomial regression model. Karlin and Studden (1966a) found explicit solutions for three types of efficiency functions. We introduce two "new" functions to model the heteroscedastic structure, for which the D-optimal designs can also be found explicitly. The optimal designs have equal masses at the roots of generalized Bessel polynomials and Jacobi-polynomials with complex parameters. It is also demonstrated that there exist no other efficiency functions such that the supporting polynomial of the D-optimal design satisfies a generalized Rodrigues' formula.

AMS 1991 Subject Classifications: Primary 62K05; Secondary 34L40
Key words and phrases: D-optimal design, weighted polynomial regression, Rodrigues' formula, Bessel polynomials, Schrödinger equation

1 Introduction

Consider the weighted polynomial regression model of degree p

$$(1.1)\qquad \begin{aligned} E[Y|x] &= \sum_{j=0}^{p} \beta_j x^j \\ V[Y|x] &= \frac{\sigma^2}{\lambda(x)} \,, \end{aligned}$$

where λ denotes a positive efficiency function, and the explanatory variable x is taken from the design space $\mathcal{X} \subset \mathbb{R}$. An approximate design ξ is a probability measure with finite support on the design space $\mathcal{X}$, and the Fisher information matrix for the parameter $\beta = (\beta_0, \dots, \beta_p)^T$ is given by the matrix

$$(1.2)\qquad M(\xi) = \int_{\mathcal{X}} \lambda(x) f(x) f^T(x) d\xi(x),$$

where $f(x) = (1, x, \dots, x^p)^T$ denotes the vector of monomials up to the order p [see Fedorov (1972), Kiefer (1974), Silvey (1980) or Pukelsheim (1993)]. A D-optimal design maximizes the determinant of the Fisher information matrix. In their pioneering work Hoel (1958) and Karlin and Studden (1966a) proved that the D-optimal designs for the efficiency functions

$$(1.3)\qquad \lambda(x) = 1; \quad \mathcal{X} = [-1, 1]$$

$$(1.4)\qquad \lambda(x) = (1-x)^{\alpha+1}(1+x)^{\beta+1}; \quad \mathcal{X} = (-1, 1) \quad (\alpha, \beta > -1)$$

$$(1.5)\qquad \lambda(x) = \exp(-x); \quad \mathcal{X} = [0, \infty)$$

$$(1.6)\qquad \lambda(x) = x^{\alpha+1}\exp(-x); \quad \mathcal{X} = (0, \infty) \quad (\alpha > -1)$$

$$(1.7)\qquad \lambda(x) = \exp(-x^2); \quad \mathcal{X} = (-\infty, \infty)$$

have equal masses at the roots of classical orthogonal polynomials [see Karlin and Studden (1966b) or Fedorov (1972) for more details]. In the following period numerous authors have worked on generalizations of these results motivated by different aspects [see Antille (1977), Huang, Chang and Wong (1995), He, Studden and Sun (1996), Chang and Lin (1997), Imhoff, Kraft and Schaefer (1998), Ortiz and Rodrigues (1998) or Dette, Haines and Imhoff (1999) among many others]. Most authors derive a differential equation for the supporting polynomial of the D-optimal design, which induces a finite dimensional eigenvalue problem. The components of the eigenvector corresponding to the minimal eigenvalue in this problem give the coefficients of the supporting polynomial. In such cases the D-opimal designs can be readily obtained numerically, but the results of Huang, Chang and Wong (1995), He, Studden and Sun (1996), Chang and Lin (1997), Ortiz

and Rodrigues (1998) and Imhoff, Krafft and Schaefer (1998) demonstrate that analytic results are in general difficult to derive.
The first purpose of this note is to give a partial explanation why only the efficiency functions of the form (1.3) - (1.7) yield D-optimal designs with support points given by the zeros of classical special functions. We use a result of Cryer (1970) to demonstrate that there are essentially five types of efficiency functions for which the solution of the D-optimal design problem is "simple" in the sense that the corresponding supporting polynomial has a representation by a generalized Rodrigues' formula. Besides the three "classical" efficiency functions specified by (1.3) - (1.7), there appear two "new" efficiency functions for which the support points of the D-optimal design problem can be specified as the zeros of classical (nonorthogonal) polynomials, namely

$$\lambda_1(x) \;=\; (1+x^2)^{\alpha+1}\exp(2\beta\arctan x); \quad \mathcal{X}=(-\infty,\infty) \tag{1.8}$$

$$\lambda_2(x) \;=\; x^{-\gamma}\exp(-\delta/x); \quad \mathcal{X}=(0,\infty)\ , \tag{1.9}$$

where $\alpha \in (-\infty, -p-1], \beta, \gamma \in \mathbb{R}, \delta \in \mathbb{R}^+$. Note that for the case $\beta = 0$ the efficiency function (1.8) has been considered by Dette, Haines and Imhoff (1999), but the general case $\beta \in \mathbb{R}\backslash\{0\}$ is not symmetric, which causes additional difficulties.
The second purpose of this note is to determine the D-optimal designs in the weighted polynomial regression model with efficiency functions (1.8) and (1.9) (for the open cases) explicitly. It will be shown that in these cases the D-optimal design puts equal masses at the $p+1$ roots of a Jacobi polynomial with complex parameters and a generalized Bessel polynomial, respectively.

2 The Rodrigues' formula

Recall that the support points of the D-optimal design for the heteroscedastic polynomial regression model with efficiency function (1.4) are given by the zeros of the Jacobi polynomial

$$P_{p+1}^{(\alpha,\beta)}(x) \tag{2.1}$$

orthogonal with respect to the measure $(1-x)^\alpha(1+x)^\beta dx$ on the interval $(-1,1)$, $(\alpha, \beta > -1)$ [see Fedorov (1972)]. Similary, the constant efficiency function (1.3) yields the zeros of the polynomial

$$(x^2-1)P_p'(x) \tag{2.2}$$

as the support points of the D-optimal design, where P_p' is the derivative of the pth Legendre polynomial orthogonal with respect to the Lebesgue measure on the interval $[-1,1]$. Note that the class of efficiency functions (1.4) essentially contains (1.3) (if $\alpha, \beta \to -1$) and it appears therefore somewhat artificial to consider these cases separately. This problem can be avoided by using a different representation for the supporting polynomials, which does <u>not</u> refer to orthogonality. To be precise, we note that the Jacobi polynomial of degree p is given by

$$P_p^{(\alpha,\beta)}(x) = \frac{(-1)^p}{2^p p!}(1-x)^{-\alpha}(1+x)^{-\beta}(\frac{d}{dx})^p\{(1-x)^{p+\alpha}(1+x)^{p+\beta}\} \tag{2.3}$$

[see Szegö (1975), p. 67]. This representation is called Rodrigues' formula and does not require the orthogonality with respect to an absolute continuous measure, which is equivalent to the condition $\alpha, \beta > -1$ for the parameters. The Legendre polynomial $P_p(x)$ is equal to $P_p^{(0,0)}(x)$ (by its definition), and formula (2.3) yields

$$(1-x^2)P_p'(x) = (1-x^2)(\frac{d}{dx})^{p+1}\{(1-x^2)^p\} = 2^{p+1}(p+1)!(-1)^{p+1}P_{p+1}^{(-1,-1)}(x)$$

for the supporting polynomial of the D-optimal design in the case of constant efficiency. Consequently, the support points of the D-optimal designs in the polynomial regression model with efficiency functions (1.3) and (1.4) are obtained as the zeros of the polynomial

$$(1-x)^{-\alpha}(1+x)^{-\beta}(\frac{d}{dx})^{p+1}\{(1-x)^{p+1+\alpha}(1+x)^{p+1+\beta}\} \tag{2.4}$$

whenever $\alpha, \beta \geq -1$. Similary, the efficiency functions (1.5) and (1.6) yield D-optimal designs supported at the zeros of the polynomial

$$x^{-\alpha}e^x(\frac{d}{dx})^{p+1}\{e^{-x}x^{p+\alpha+1}\} \tag{2.5}$$

whenever $\alpha \geq -1$. Note that for $\alpha > -1$ this polynomial is proportional to $L_{p+1}^{(\alpha)}(x)$ [see Szegö (1975), p. 101], while for $\alpha = -1$ it gives $xL_p^{(1)}(x)$ up to a constant. Finally, the support of the D-optimal design for polynomial regression with efficiency (1.7) is given by the zeros of the $(p+1)$th Hermite polynomial proportional to

$$e^{x^2}(\frac{d}{dx})^{p+1}\{e^{-x^2}\} \tag{2.6}$$

[see Szegö (1975), p. 106]. From this point of view there exist only three types of efficiency functions and the support points of the D-optimal design in the corresponding heteroscedastic polynomial regression model are given by the zeros of a polynomial, which can be represented by a generalized Rodrigues' formula [see Erdelyi, Magnus, Oberhettinger, Tricomi (1953)] of the form

$$\frac{1}{\omega(x)}(\frac{d}{dx})^{p+1}\{\omega(x)\cdot h^{p+1}(x)\} \tag{2.7}$$

where h is a given polynomial and ω an arbitrary function. The following result proved by Cryer (1970) characterizes the class of functions ω and polynomials h such the generalized Rodrigues' formula defines a polynomial of degree $p+1$ for each $p = -1, 0, 1, \ldots$.

Theorem [Cryer (1970)]. *If the generalized Rodrigues' formula (2.7) defines a polynomial of degreee $p+1$ for $p = -1, 0, 1, \ldots$, then the function ω and the polynomial h are of the following type (modulo affine transformations)}*

$$w(x) = e^{-x^2}, \quad h(x) = 1 \tag{2.8}$$

(2.9) $$w(x) = x^a e^{-x}\ , \quad h(x) = x$$

(2.10) $$w(x) = (1-x)^a(1+x)^b\ , \quad h(x) = (1-x^2)$$

(2.11) $$w(x) = x^{-a}e^{-b/x}\ , \quad h(x) = x^2$$

(2.12) $$w(x) = (1+x^2)^a e^{b \arctan x}\ , \quad h(x) = (1+x^2).$$

Note that (2.8) - (2.10) correspond to the classical cases (2.4) - (2.6) with efficiency functions given by (1.3) - (1.7). However, there are two new cases, which have not been considered so far and correspond to the efficiency functions in (1.8) and (1.9). The corresponding D-optimal design problems will be discussed in the following section.

3 D-optimal design problems for weighted polynomial regression – two new results

Consider the polynomial regression model (1.1) with efficiency function (1.8). In order to guarantee the existence of an optimal design on the design space $\mathbb{R}$, the induced design space

$$\left\{(1, x, \dots, x^p)^T \lambda(x) \ \mid \ x \in \mathbb{R}\right\}$$

has to be bounded, which requires $\alpha \le -p-1$ in (1.8). For such cases the D-optimal design can also be described by the roots of Jacobi polynomials using complex parameters.

Theorem 3.1. *The D-optimal design for the weighted polynomial regression model with efficiency function*

$$\lambda_1(x) = (1+x^2)^{\alpha+1} \exp(2\beta \arctan x)$$

and $\alpha \le -p-1$ puts equal masses at the zeros of the Jacobi polynomial

$$P_{p+1}^{(\alpha+i\beta, \alpha-i\beta)}(xi)$$

defined by (2.4).

Proof. In the case $\beta = 0$ the result is reduced to Theorem 3.1 of Dette, Haines and Imhof (1999) and therefore we restrict ourselves to the case $\beta \ne 0$ throughout this proof. Careful inspection of the directional derivative shows that for $\alpha < -p-1$ the D-optimal design has $p+1$ support points, and a standard argument shows that the optimal weights at these points have to be equal. The determinant of a design with equal weights at the $p+1$ points $x_0, \dots, x_p$ is proportional to

$$\prod_{i=0}^{p} \lambda_1(x_i) \prod_{0 \le i < k \le p} (x_i - x_k)^2.$$

Taking partial derivatives and using the same arguments as in Karlin and Studden (1996b) we obtain the differential equation

$$(3.1)\qquad (1+x^2)y'' + 2[\beta + (\alpha+1)x]y' - (p+1)[p+2(\alpha+1)]y = 0$$

for the supporting polynomial $g(x) = \prod_{i=0}^{p}(x-x_i)$. This gives for the function $\tilde{y}(x) = y(ix)$ the differential equation

$$(3.2)\qquad (1-x^2)\tilde{y}'' + 2[\beta i - (\alpha+1)x]\tilde{y}' + (p+1)[p+2(\alpha+1)]\tilde{y} = 0.$$

It is well known from the theory of hypergeometric functions that a fundamental set of solutions of the differential equation (3.2) is given by

$$\begin{aligned} \tilde{y}_1(x) &= F\Big(-p-1;\; p+2\alpha+2;\; \alpha-\beta i+1;\; \frac{1-x}{2}\Big) \\ \tilde{y}_2(x) &= (1-x)^{-\alpha+\beta i}\cdot F\Big(-p-1-\alpha+\beta i;\; p+2+\alpha+\beta i;\; 1-\alpha+\beta i;\; \frac{1-x}{2}\Big)\ , \end{aligned}$$

where

$$F(a;\; b;\; c;\; x) \;=\; 1+\sum_{\nu=1}^{\infty}\frac{a(a+1)\cdots(a+\nu-1)}{1\cdot 2\cdots\nu}\,\frac{b\cdot(b+1)\cdots(b+\nu-1)}{c\cdot(c+1)\cdots(c+\nu-1)}\cdot x^{\nu}$$

denotes the hypergeometric series [see Whittaker and Watson (1935), p. 283-286 or Andrews, Askey and Roy (1999), p. 75-81]. For $\beta \neq 0$ the function $\tilde{y}_2$ is not a polynomial, while $\tilde{y}_1$ is proportional to the Jacobi polynomial $P_{p+1}^{(\alpha-i\beta,\alpha+i\beta)}(x)$ [see Szegö (1975), p. 62], which yields for the supporting polynomial

$$\begin{aligned} g(x) &= \prod_{j=0}^{p}(x-x_j) = \tilde{y}(-ix) = c\cdot P_{p+1}^{(\alpha-i\beta,\alpha+i\beta)}(-ix) \\ &= c(-1)^{p+1}P_{p+1}^{(\alpha+i\beta,\alpha-i\beta)}(ix), \end{aligned}$$

where the last equality follows from the symmetry property of the Jacobi polynomials [see Szegö (1975), p. 59], and the constant c is defined such that the leading coefficient of the right hand side is equal to one. This proves the assertion for $\alpha < -p-1$ and the remaining case $\alpha = -p-1$ follows by continuity.

□

Remark 3.2. There is an intuitive explanation of the result of Theorem 3.1. To be precise, observe that

$$\arctan z = \frac{-i}{2}\log\frac{1+iz}{1-iz}\ ,$$

which gives for the efficiency function

$$\lambda_1(z) = (1+z^2)^{\alpha+1}\left(\frac{1+iz}{1-iz}\right)^{-\beta i} \;=\; (1-iz)^{\alpha+\beta i+1}(1+iz)^{\alpha-\beta i+1}\ ,$$

and a naive generalization of the classical cases (1.3) and (1.4) yields the assertion of Theorem 3.1.

We will conclude this section giving the corresponding statement for the efficiency function (1.9).

Theorem 3.3. *The D-optimal design for the weighted polynomial regression model with efficiency function*

$$\lambda_2(x) = x^{-\gamma} \exp(-\delta/x)$$

and $\delta > 0$ puts equal masses at the roots of the generalized Bessel polynomial

$$Y_{p+1}(x, -\gamma, \delta) = \sum_{k=0}^{p+1} \binom{p+1}{k} (p+k-\gamma+1)^{(k)} (\frac{x}{\delta})^k \tag{3.3}$$

where $z^{(0)} := 1$ and $z^{(k)} := z(z-1)\dots(z-k+1)$ if $k \geq 1$.

Proof. The same arguments as given in the proof of Theorem 3.1 show that the D-optimal design is supported at $p+1$ points $x_0, \dots, x_p$, and that the supporting polynomial $g(x) = \prod_{j=0}^{p}(x - x_j)$ satisfies the second order differential equation

$$x^2 y'' + (\beta - \gamma x) y' - (p+1)(p-\gamma) y = 0. \tag{3.4}$$

It now follows from the results of Krall and Frink (1949) that $g(x)$ is proportional to the generalized Bessel polynomial defined in (3.3).

□

Acknowledgements. The authors are thankful to the 3rd-cycle romand de statistique et probabilites appliquees, which provides the opportunity to start this joint work. The authors are also grateful to Isolde Gottschlich who typed most parts of this paper with considerable technical expertise. Parts of this paper were written, during the second author's visit at Purdue University and this author would like to thank the Department of Statistics for its hospitality. The financial support of the Deutsche Forschungsgemeinschaft (SFB 475, Reduction of complexity in multivariate data structures) is gratefully acknowledged.

References

G.E. Andrews, R. Askey, R. Roy (1999). Special functions. Encyclopedia of Mathematics and its Applications. Cambridge, University Press.

G. Antille (1977). Plans d'expérience. Measures et plans d'expérience D-optimaux pour les polynòmes de régression. C.R. Acad. Sci. Paris 285, 577-580.

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

C.W. Cryer (1970). Rodrigues' formula and the classical orthogonal polynomials. Bolletino della Unione Mathematica Italiana, Ser. IV, Anno III, 1-11.

H. Dette, L.M. Haines, L. Imhoff (1999). Optimal designs for rational models and weighted polynomial regression. Ann. Statist. 27, 1272-1293.

A. Erdélyi, W. Magnus, F. Oberhettinger, F.G. Tricomi (1953). Higher transcendental functions, Vol. II. McGraw Hill, NY.

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

Z. He, W.J. Studden, D. Sun (1996). Optimal designs for rational models. Ann. Statist. 24, 2128-2147.

P.G. Hoel (1958). Efficiency problems in polynomial estimation. Ann. Math. Statist. 29, 1134-1145.

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

L. Imhoff, O. Krafft, M. Schaefer (1998). D-optimal designs for polynomial regression with weight function $\omega(x) = x/(1+x)$. Statist. Sinica 8, 1271-1274.

S. Karlin, W.J. Studden (1966a). Optimal experimental designs. Ann. Math. Statist. 37, 783-815.

S. Karlin, W.J. Studden (1966b). Tchebycheff systems: with applications in analysis and statistics. Wiley, New York.

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

H.L. Krall, O. Frink (1949). A new class of orthogonal polynomials: the Bessel polynomials. Trans. Amer. Math. Soc. 65, 100-115.

I. Oritz, C. Rodrigues (1998). D-optimal designs for weighted polynomial regression without any initial terms. In MODA 5. *Advances in Model-Oriented Data Analysis and Experimental Design* (eds. A.C. Atkinson, L. Pronzato, H.P. Wynn), 67-74.

F. Pukelsheim (1993). Optimal design of experiments. Wiley, New York.

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

G. Szegö (1975). Orthogonal polynomials. Amer. Math. Soc. Colloqu. Publ. 23, Providence, RI.

E.T. Whittaker, G.N. Watson (1935). A course of modern analysis. Cambridge.