Constrained D- and D_1-optimal designs for polynomial regression

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

Tobias Franke
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany
email: tobias.franke@ruhr-uni-bochum.de

July 20, 2000

Abstract

In the common polynomial regression model of degree m we consider the problem of determining the D- and D_1-optimal designs subject to certain constraints for the D_1-efficiencies in the models of degree $m-j, m-j+1, \ldots, m+k$ ($m > j \geq 0, k \geq 0$ given). We present a complete solution of these problems, which on the one hand allow a fast computation of the constrained optimal designs and on the other hand give an answer to the question of the existence of a design satisfying all constraints. Our approach is based on a combination of general equivalence theory with the theory of canonical moments. In the case of equal bounds for the D_1-efficiencies the constrained optimal designs can be found explicitly by an application of recent results for associated orthogonal polynomials.

AMS Subject classification: 62K05, 33C45
Keywords and Phrases: Constrained optimal designs, polynomial regression, D- and D_1-optimal designs, associated orthogonal polynomials

1 Introduction

In fitting a parametric regression model there are usually several objectives which should be addressed by the design of the experiment. Model adequacy could be a serious problem [see e.g. Box and Draper (1959)] and usually designs are desirable which are on the one hand efficient for discriminating between several competing models and have on the other hand good properties for the estimation of the parameters in the identified model. Multiple objectives cannot be easily characterized by standard optimality criteria as proposed in Kiefer (1974).
There are essentially two ways for the construction of design criteria which incorporate different purposes of the experimenter. One approach is the construction of a new optimality criterion

by averaging several competitive design criteria. This is called a compound (or weighted) optimal design problem [see Läuter (1974), Cook and Nachtsheim (1982) or Dette (1990)]. Alternatively one could try to maximize one primary optimality criterion subject to constraints for specific minimum efficiencies of other criteria. This is called a constrained optimal design problem [see Stigler (1971), Studden (1982b), Lee (1988 a,b)]. Cook and Wong (1994) showed the equivalence between compound and constrained optimal designs in the case of two design criteria [see also Dette (1995a) and Clyde and Chaloner (1996) for more general formulations of these results]. Roughly speaking, a solution of a compound optimal design problem is also a solution of the constrained optimal design problem, if the constraints are defined appropriately. Conversely, an appropriate definition of the weights yields always optimality of a constrained optimal design with respect to the compound criterion [see Dette (1995a)].

Although these results are interesting from a theoretical point of view they are not too useful for determining constrained optimal designs in practice. In most cases these designs have to be found numerically, and the corresponding algorithms only work with a few constraints. A particular difficulty in such calculations is the determination of a starting design, because the question of the existence of at least one design satisfying all constraints has in general no clear answer. Moreover, to the knowledge of the authors no explicit solutions are available for constrained optimal design problems with more than one constraint. One reason for these difficulties is that the corresponding equivalence theorems contain certain Lagrange multipliers which in general are not uniquely determined [see e.g. Pukelsheim (1993), Section 11.19 and 11.20].

It is the purpose of the present paper to provide additional insight in the complicated structure of these problems by deriving explicit solutions for two constrained optimal design problems which appear in polynomial regression models. Our first criterion (called constrained D_1-D_1-criterion) is motivated by the identification of the appropriate degree of the polynomial. Here the constrained optimal design maximizes the power of the test for the highest coefficient in a model of degree m subject to the constraints that the design yields efficient tests for the highest coefficients in the models of degree $m-j, m-j+1, \ldots, m-1, m+1, \ldots, m+k$ ($m > j \geq 0, k \geq 0$ given). Similary, our second criterion (termed constrained D-D_1-criterion) determines the D-optimal design for the model of degree m in the class of all designs which guarantee given efficiencies for testing the highest coefficients in the models of degree $m-j, m-j+1, \ldots, m+k-1, m+k$. Our approach is based on a combination of general equivalence theory [see Pukelsheim (1993)] with the theory of canonical moments which was introduced by Skibinsky (1967) and applied by Studden (1980, 1982a, 1982b, 1989) for determining optimal designs in polynomial regression models. This enables us to identify the Lagrange multipliers in the corresponding equivalence theorem explicitly and to characterize the constrained optimal design by a system of (nonlinear) equations for its canonical moments. Moreover, for special choices of the constraints (e.g. equal constraints for all efficiencies) the corresponding designs can be characterized by linear combinations of associated ultraspherical polynomials [see Grosjean (1986) or Lasser (1994)].

The paper will be organized as follows. In Section 2 we introduce the constrained optimality criteria in the context of polynomial regression models. The D_1-D_1-constrained optimal design problem is solved in Section 3, while Section 4 states the corresponding results for the D-D_1-constrained optimality criterion. Finally, some of the more technical proofs are deferred to the appendix in Section 5.

2 Two constrained design criteria for polynomial regression models

Consider the common polynomial regression of degree m

$$Y = \vartheta_m^T f_m(x) + \varepsilon = \sum_{j=0}^{m} \vartheta_{m,j} x^j + \varepsilon \tag{2.1}$$

where ε is a random error with mean 0 and constant variance, $f_m(x) = (1, x, \ldots, x^m)^T$ denotes the vector of regression functions, $\vartheta_m = (\vartheta_{m,0}, \ldots, \vartheta_{m,m})^T$ is the vector of parameters and the explanatory variable is taken from a compact interval, say $\mathcal{X}$. For an approximate design ξ, which is a probability measure with finite support on the design space $\mathcal{X}$, the Fisher information matrix for the parameter ϑ_m can be expressed as

$$M_m(\xi) = \int_{\mathcal{X}} f_m(x) f_m^T(x) d\xi(x). \tag{2.2}$$

An optimal design maximizes an appropriate information function of the Fisher information matrix $M_m(\xi)$ [see Pukelsheim (1993)] and there are numerous criteria which can be used for the construction of efficient designs. In this paper we will concentrate on the D-criterion

$$\Phi_m(\xi) := |M_m(\xi)|^{1/(m+1)} \to \max_{\xi} \tag{2.3}$$

and the D_1-criterion

$$\Psi_m(\xi) := (e_m^T M_m^{-1}(\xi) e_m)^{-1} = \frac{|M_m(\xi)|}{|M_{m-1}(\xi)|} \to \max_{\xi} \tag{2.4}$$

where $e_m = (0, \ldots, 0, 1)^T \in \mathbb{R}^{m+1}$ denotes the $(m+1)$th unit vector. A D-optimal design minimizes the volume of the ellipsoid of concentration for the unknown parameter $\vartheta_m \in \mathbb{R}^{m+1}$ while a D_1-optimal design maximizes the power of the test for the hypothesis $H_0 : \vartheta_{m,m} = 0$ in the polynomial regression of degree m. The D- and D_1-optimal designs for the polynomial regression model of degree m have been explicitly found by Hoel (1958) [see also Guest (1958)] and Kiefer and Wolfowitz (1959), respectively.
Note that the two optimality criteria in (2.3) and (2.4) require the specification of a „correct" degree of the regression. In practice this is rarely available and a design for a degree m model will typically be used to test the terms in the model for significance or to test the lack of fit of higher and lower order polynomials. If, for example, the experimenter has some prior information that a polynomial of degree m adequately describes the data but wants to use his design to test for polynomials of degree $m-j, m-j+1, \ldots, m+k-1, m+k$ for given $j, k \in \mathbb{N}_0, j < m$, the following two constrained optimality criteria might be appropriate

$$\begin{aligned} &\text{maximize } \Psi_m(\xi) \text{ subject to} \\ &\text{eff}_l^{D_1}(\xi) \geq c_l \quad \text{for } l = m-j, \ldots, m-1, m+1, \ldots, m+k \end{aligned} \tag{2.5}$$

$$\begin{aligned} &\text{maximize } \Phi_m(\xi) \text{ subject to} \\ &\text{eff}_l^{D_1}(\xi) \geq c_l \quad \text{for } l = m-j, \ldots, m+k. \end{aligned} \tag{2.6}$$

Here $\mathrm{eff}_l^{D_1}(\xi)$ denotes the efficiency of the design ξ for testing the highest coefficient in the polynomial regression of degree l, that is

$$\mathrm{eff}_l^{D_1}(\xi) = \frac{\Psi_l(\xi)}{\max\limits_{\eta} \Psi_l(\eta)}, \tag{2.7}$$

and $c_{m-j}, \ldots, c_{m+k} \in (0,1)$ denote given numbers specifying the guaranteed efficiencies of the design ξ for testing the highest coefficients in the models of degree $m-j, \ldots, m+k$. The criterion (2.5) could be used if the primary interest of the experiment is the identification of the appropriate degree of the regression and there is some preference for the model of degree $m-1$ or m. Similary the constrained optimality criterion (2.6) is useful, if a model of degree m seems to be appropriate but there is a possibility of a higher or lower order regression. In this case the maximization of the determinant in (2.6) will yield a good design for estimating the parameters in the model of degree m which has reasonable efficiencies for testing the lack of fit of the polynomials of degree $m-j, \ldots, m+k$.
It follows from standard arguments in design theory that for the polynomial regression model the constrained optimization problems (2.5) and (2.6) are not changed under an affine transformation of the design space $\mathcal{X}$ and we may assume without loss of generality $\mathcal{X} = [-1, 1]$. Moreover, the strict concavity of the criteria in (2.3) and (2.4) implies that a constrained optimal design [with respect to (2.5) or (2.6)] on a symmetric design space must be symmetric. A further important tool for determining optimal designs for polynomial regression is the theory of canonical moments which was introduced by Skibinsky (1967) and applied by Studden (1980, 1982a, 1982b) in this context [see also Lau (1983, 1988), Skibinsky (1986) and the recent monograph of Dette and Studden (1997)]. Roughly speaking every probability measure on the interval $[-1, 1]$ is uniquely determined by a sequence $(p_1, p_2, \ldots)$ whose elements vary independently in the interval $[0, 1]$. For a given probability measure on the interval $[-1, 1]$ the element p_j of the corresponding sequence is called the jth canonical moment of ξ. If j is the first index for which $p_j \in \{0, 1\}$, then the sequence of canonical moments terminates at p_j, the measure is supported at a finite number of points and can be determined by evaluating certain orthogonal polynomials [see Skibinsky (1986) or Lau (1988)]. Moreover, a measure ξ on the interval $[-1, 1]$ is symmetric if and only if all canonical moments of odd order are equal 1/2 and for a symmetric measure we obtain for the determinant of the information matrix

$$|M_m(\xi)| = \prod_{j=1}^{m} (q_{2j-2} p_{2j})^{m-j+1}. \tag{2.8}$$

This gives for the element in the position $(m+1, m+1)$ of the matrix $M_m^{-1}(\xi)$

$$(e_m^T M_m^{-1}(\xi) e_m)^{-1} = \prod_{j=1}^{m} q_{2j-2} p_{2j}, \tag{2.9}$$

where $p_2, p_4, \ldots$ denote the canonical moments (of even order) of the design ξ and $q_j = 1 - p_j \quad (j \geq 1)$. Observing these identities we can easily identify the canonical moments of the D_1-optimal design in the polynomial regression model of degree m, i.e.

$$p_j = \frac{1}{2}, \qquad j = 1, \ldots, 2m-1; \quad p_{2m} = 1,$$

which implies for the D_1-efficiency of a symmetric design ξ

$$\text{(2.10)} \qquad \mathrm{eff}_m^{D_1}(\xi) = 2^{2m-2} \prod_{i=1}^{m} q_{2i-2} p_{2i}.$$

Similary, the D-optimal design in the polynomial regression of degree m is determined by its canonical moments

$$\text{(2.11)} \qquad p_{2l} = \frac{m-l+1}{2(m-l)+1}, \qquad p_{2l-1} = \frac{1}{2}, \qquad l = 1, \dots, m$$

where the canonical moments of even order are obtained by maximizing (2.8) and the canonical moments of odd order are all equal to $1/2$ by the symmetry of the D-optimal design.
It will be demonstrated in Section 3 and 4 that the constrained optimal designs (with respect to the criteria (2.5) and (2.6) can be described explicitly by a system of (nonlinear) equations for their canonical moments. In other words, the solution of this system yields the canonical moments of the constrained optimal design and the identification of the measure corresponding to the ,,optimal" canonical moments can then be performed by standard methods [see Dette and Studden (1997), Section 3].

3 D_1-D_1-constrained optimal designs

An important tool for the verification of optimality and the numerical construction of optimal designs are equivalence theorems [see Kiefer (1974)]. In the context of constrained optimal design criteria these characterizations contain several Lagrange multipliers which are not necessarily uniquely determined. Although these unknown quantities in the characterizing inequality make a direct application of the equivalence theorem impossible, we will nevertheless use this type of characterization as one main ingredient for the solution of the constrained optimal design problem (2.5).

Theorem 3.1. *A design ξ is a solution of the constrained optimal design problem (2.5) if and only if ξ satisfies the constraints and there exist nonnegative numbers α_l ($l = m-j, \dots, m-1, m+1, \dots, m+k$), $\alpha_{m+k} > 0$, with*

$$\text{(3.1)} \qquad \alpha_l \, \mathrm{eff}_l^{D_1}(\xi) \;=\; \alpha_l c_l$$

$l = m-j, \dots, m-1, m+1, \dots, m+k$

$$\text{(3.2)} \qquad \alpha_m \;=\; 1 - \sum_{l=m-j}^{m-1} \alpha_l - \sum_{l=m+1}^{m+k} \alpha_l \geq 0$$

such that the inequality

$$\text{(3.3)} \qquad \sum_{l=m-j}^{m+k} \alpha_l \frac{(e_l^T M_l^{-1}(\xi) f_l(x))^2}{e_l^T M_l^{-1}(\xi) e_l} \leq 1$$

holds for all $x \in \mathcal{X} = [-1, 1]$.

Proof. Note that $c_{m+k} > 0$ implies $|M_{m+k}(\eta)| > 0$ for any design satisfying the constraints in (2.5). If there exists a design η with

$$\text{eff}_l^{D_1}(\eta) > c_l \quad l = m-j, \ldots, m-1, m+1, \ldots, m+k \tag{3.4}$$

the assertion of the theorem is a direct consequence of a general equivalence theorem for constrained optimal designs [see Pukelsheim (1993), Section 11.19 and 11.20]. In this case we have additionally $\alpha_m > 0$.

In the remaining case any design which satisfies the constraints in (2.5) (now with at least one equality) also maximizes

$$\min\{\frac{\text{eff}_l^{D_1}(\xi)}{c_l} \mid l = m-j, \ldots, m-1, m+1, \ldots, m+k\}$$

in the class of all designs (such that $|M_{m+k}(\xi)| \neq 0$) with optimal value equal to one. Now the results of Dette (1995b) [Theorem 2.1 in this paper with $n = m+k$, $\vartheta_l^2 = \{\max_\eta (e_l^T M_l^{-1}(\eta) e_l)^{-1} c_l\}^{-1}$ if $l = m-j, \ldots, m-1, m+1, \ldots, m+k$, $\vartheta_l = 0$ if $l = 1, \ldots, m-j-1, m$] show that there is exactly one design maximizing the minimum of weighted D_1-efficiencies which is obviously optimal with respect to the constrained criterion (2.5) and satisfies (3.1) - (3.3) with $\alpha_m = 0$. □

Theorem 3.2. *The solution of the constrained optimal design problem (2.5) is unique. A design ξ^* is a solution of the constrained optimal design problem (2.5) if and only if its canonical moments of odd order are given by $p_{2j-1} = \frac{1}{2}$ $(j = 1, \ldots, m+k)$, $p_{2m+2k} = 1$, and the canonical moments $(p_2, \ldots, p_{2m+2k-2})$ of even order are the solution of the system of equations*

$$p_{2l} = \frac{1}{2}, \tag{3.5}$$

$l = 1, \ldots, m-j-1$

$$p_{2l} = \max\Big\{\frac{c_l}{2^{2(l-m+j)} \prod_{i=m-j}^{l-1} p_{2i}(1-p_{2i})}, \frac{1}{2}\Big\}, \tag{3.6}$$

$l = m-j, \ldots, m-1,$

$$p_{2l} = \max\Big\{1 - \frac{c_{m+k}}{2^{2(m+k-l)} c_l \prod_{i=l+1}^{m+k-1} p_{2i}(1-p_{2i})}, \frac{1}{2}\Big\}, \tag{3.7}$$

$l = m+k-1, m+k-2, \ldots m+1,$

$$p_{2m} = \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{c_{m+k}}{2^{2(j+k)} \prod_{l=m-j, l\neq m}^{m+k-1} p_{2l}(1-p_{2l})}} \quad \textit{if } k > 0. \tag{3.8}$$

Proof. By the discussion of Section 2 the constrained optimal design must be symmetric and we obtain $p_{2j-1} = 1/2$; $\quad j = 1, \ldots, m+k$. By Theorem 3.1 of this paper and Theorem 6.3.2 in Dette and Studden (1997) (for $p = 0$) a design ξ^* is a solution of the constrained optimal design problem (2.5) if and only if there exists a prior $(\beta_1, \ldots, \beta_{m+k})$ for the class of polynomials of degree $1, 2, \ldots, m+k$ with $(\beta_l \geq 0, \sum_{l=1}^{m+k} \beta_l = 1, \beta_{m+k} > 0)$

$$\beta_l = 0 \quad l = 1, \ldots, m-j-1 \tag{3.9}$$

$$\beta_l \mathrm{eff}_l^{D_1}(\xi^*) = \beta_l c_l \quad l = m-j, \ldots, m-1, m+1, \ldots, m+k \ , \tag{3.10}$$

such that ξ^* maximizes the geometric mean

$$\prod_{l=1}^{m+k} (\mathrm{eff}_l^{D_1}(\eta))^{\beta_l}$$

over the class of all designs with $|M_{m+k}(\eta)| \neq 0$. Now Theorem 6.2.3 in Dette and Studden (1997) implies $p_{2m+2k} = 1$ and Theorem 6.2.6 in the same reference expresses the weights β_l in terms of the canonical moments of the optimal design ξ^*, that is

$$\beta_l = \prod_{j=1}^{l-1} \frac{q_{2j}}{p_{2j}} (1 - \frac{q_{2l}}{p_{2l}}) \quad l = 1, \ldots, m+k \tag{3.11}$$

(note that by symmetry all canonical moments of odd order are equal 1/2). From (3.9) and (3.11) we thus obtain (3.5), i.e. $p_{2l} = 1/2; l = 1, \ldots, m-j-1$. Now (2.10) gives

$$\mathrm{eff}_{m-j}^{D_1}(\xi^*) = 2^{2m-2j-2} \prod_{i=1}^{m-j} q_{2i-2} p_{2i} = p_{2m-2j}$$

and using (3.10) for $l = m-j$ yields for the canonical moment of order $2m-2j$ either $p_{2m-2j} = 1/2$ (equivalently $\beta_{m-j} = 0$) or $p_{2m-2j} = c_{m-j}$. Observing the inequalities $\mathrm{eff}_{m-j}^{D_1}(\xi^*) \geq c_{m-j}, \beta_{m-j} \geq 0$ gives

$$p_{2m-2j} = \max\{c_{m-j}, \frac{1}{2}\} =: \alpha$$

(note that $p_{2m-2j} q_{2m-2j}$ appears in all efficiencies in the models of degree $l > m-j$ and consequently increasing p_{2m-2j} in the interval $[\alpha, \frac{1}{2}]$ will yield smaller efficiencies in the models of higher degree).
In the next step we obtain for $l = m-j+1$

$$\mathrm{eff}_{m-j+1}^{D_1}(\xi^*) = 2^{2(m-j)} \prod_{i=1}^{m-j+1} q_{2i-2} p_{2i} = 2^2 q_{2m-2j} p_{2m-2j} p_{2m-2j+2}$$

which implies (by the same reasoning)

$$p_{2m-2j+2} = \max\{\frac{c_{m-j+1}}{2^2 q_{2m-2j} p_{2m-2j}}, \frac{1}{2}\}.$$

Repeating these arguments for $l = m-j, \ldots, m-1$ yields the canonical moments in (3.6). For a proof of the identity (3.7) we note that $\beta_{m+k} > 0$ implies (by an application of (3.10) for $l = m+k$)

$$\text{(3.12)} \qquad \mathrm{eff}^{D_1}_{m+k}(\xi^*) = c_{m+k}.$$

A further application of (3.10) (for $l = m + k - 1$) and (2.10) now yields

$$\frac{\beta_{m+k-1}}{4q_{2m+2k-2}} = \frac{\beta_{m+k-1}\mathrm{eff}^{D_1}_{m+k-1}(\xi^*)}{\mathrm{eff}^{D_1}_{m+k}(\xi^*)} = \beta_{m+k-1}\frac{c_{m+k-1}}{c_{m+k}}$$

which can be solved with respect to $p_{2m+2k-2}$. This yields either $p_{2m+2k-2} = \frac{1}{2}$ (equivalently $\beta_{m+k-1} = 0$) or $p_{2m+2k-2} = 1 - c_{m+k}/(2^2c_{m+k-1})$. Observing $\beta_{m+k-1} \geq 0$ and that the maximization of $p_{2m+2k-2}$ makes $\mathrm{eff}^{D_1}_{m+k-1}(\xi^*)$ as large as possible (without affecting the efficiencies in the models of degree yields $l \leq m + k - 2$) yields

$$p_{2m+2k-2} = \max\{1 - \frac{c_{m+k}}{2^2c_{m+k-1}}, \frac{1}{2}\}.$$

Repeating these steps for $l = m + k - 2, \ldots, m + 1$ gives the identities in (3.7). Finally, if $k > 0$ (3.12) is a quadratic equation with respect to p_{2m}, where p_{2m} should be as large as possible [because $\mathrm{eff}^{D_1}_m(\xi)$ has to be maximized]. This implies (3.8) and proves that the canonical moments (of even order) of the solution of the constrained optimal design problem satisfy the equations (3.5) – (3.8). Conversely, these arguments also show that a design with canonical moments specified in Theorem 3.2 satisfies the conditions (3.1) – (3.3) of Theorem 3.1, which is equivalent to its optimality with respect to the constrained optimality criterion (2.5). This proves the main part of the assertion of Theorem 3.2.

The remaining statement regarding the uniqueness is shown as follows. It follows from the previous discussion for any optimal design $p_{2m+2k} = 1$. This implies [observing (3.10) and (3.11)] $\mathrm{eff}^{D_1}_{m+k}(\xi) = c_{m+k}$ for any design ξ which is optimal with respect to the constrained optimality criterion (2.5). Assume that $\xi^{(1)}$ and $\xi^{(2)}$ were optimal designs with respect to the criterion (2.5) with corresponding canonical moments $p^{(1)}_j, p^{(2)}_j$, respectively ($p^{(1)}_{2m+2k} = p^{(2)}_{2m+2k} = 1$), then the concavity of the function

$$\log \mathrm{eff}^{D_1}_l(\xi) = \sum_{j=1}^{l-1} \log p_{2j}(1 - p_{2j}) + \log p_{2l} + (2l - 2)\log 2$$

($l = 1, \ldots, m + k - 1$) and the strict concavity of the function

$$\log \mathrm{eff}^{D_1}_{m+k}(\xi) = \sum_{j=1}^{m+k-1} \log p_{2j}(1 - p_{2j}) + (2m + 2k - 2)\log 2$$

on the cube $(0, 1)^{m+k-1}$ imply that the design ξ^* corresponding to the canonical moment $p^*_j = (p^{(1)}_j + p^{(2)}_j)/2$ is also optimal with respect to the constrained optimality criterion (2.5) and additionally satisfies $\mathrm{eff}^{D_1}_{m+k}(\xi^*) > c_{m+k}$. But this contradicts $p^*_{2m+2k} = 1$ (which implies $\beta^*_{m+k} > 0$) and (3.10) proving the uniqueness of the solution of the constrained optimal design problem. □

Note that Theorem 3.2 also answers the question of the existence of designs satisfying all constraints in (2.5). The solution of the constrained optimal design problem is either unique or there does not exist any design satisfying all constraints. In the first case all canonical moments

$p_{2m-2j}, \ldots, p_{2m+2k-2}$ specified by (3.6), (3.7) and (3.8) are located in the interval $(0,1)$. If there exists no design satisfying all constraints in (2.5) either the equations (3.6) define a quantity outside the interval $(0,1)$ or p_{2m} defined in (3.8) is not real. In this case the quantities p_j defined in Theorem 3.2 are no canonical moments (as introduced in Section 2) and do not correspond to a design on the interval $[-1,1]$. Consequently there is no solution of the constrained optimal design problem (2.5). We will illustrate both situations in the following example.

Example 3.3. Consider the case $m=2, k=1, j=1$, where the constrained optimal design problem (2.5) simplifies to

$$\text{(3.13)} \qquad \begin{aligned} \text{maximize} \;\; & \mathrm{eff}_2^{D_1}(\xi) \text{ subject to} \\ & \mathrm{eff}_1^{D_1}(\xi) \geq c_1 \\ & \mathrm{eff}_3^{D_1}(\xi) \geq c_3. \end{aligned}$$

In other words we are interested in a good design for testing the coefficient of the quadratic term in a polynomial of degree 2 with guaranteed efficiencies for testing the highest coefficient in the linear and cubic model. From Theorem 3.2 we obtain $p_{2j-1} = 1/2 \quad j = 1,2,3, p_6 = 1$ and

$$\begin{aligned} p_2 &= \max\{\frac{1}{2}, c_1\} \\ p_4 &= \begin{cases} \frac{1}{2}(1+\sqrt{1-c_3}) & \text{if } c_1 \leq \frac{1}{2} \\ \frac{1}{2}(1+\sqrt{1-\frac{c_3}{4c_1(1-c_1)}}) & \text{if } c_1 > \frac{1}{2} \end{cases} \end{aligned}$$

(note that the equation (3.7) does not appear in this case). If $c_1 \leq 1/2$ there always exists a solution ξ^* of the constrained optimal design problem (3.13) with efficiencies

$$\mathrm{eff}_1^{D_1}(\xi^*) = \frac{1}{2}, \;\; \mathrm{eff}_2^{D_1}(\xi^*) = \frac{1}{2}(1+\sqrt{1-c_3}), \;\; \mathrm{eff}_3^{D_1}(\xi^*) = c_3.$$

If $c_1 > 1/2$ the constrained optimal design problem (3.13) is solvable if and only if the bounds for the efficiencies satisfy

$$c_3 \leq 4c_1(1-c_1).$$

In this case the constrained D_1-D_1-optimal design ξ^* yields the D_1-efficiencies

$$\mathrm{eff}_1^{D_1}(\xi^*) = c_1, \quad \mathrm{eff}_2^{D_1}(\xi^*) = 2c_1(1-c_1)\Big(1+\sqrt{1-\frac{c_3}{4c_1(1-c_1)}}\Big), \quad \mathrm{eff}_3^{D_1}(\xi^*) = c_3.$$

In all cases the constrained optimal design puts masses $\alpha = p_2p_4/(2(1-p_2q_4))$ at the points -1 and 1 and masses $1/2-\alpha$ at the points $-\sqrt{p_2q_4}$ and $\sqrt{p_2q_4}$, respectively [see Dette and Studden (1997), p. 106].

In the remaining part of this section we will concentrate on the special but very important case that all constraints for the efficiencies in (2.5) are equal. The following result is obtained by an application of Theorem 3.2.

Corollary 3.4. *A solution of the constrained optimal design problem (2.5) with* $c_l = c \in (0,1); l = m-j,\dots,m-1,m+1,\dots,m+k$ *exists if and only if*

$$0 < c \leq \frac{j+k+1}{2(j+k)} \qquad \textit{in the case } k > 0. \tag{3.14}$$

$$0 < c < \frac{j+1}{2j} \qquad \textit{in the case } k = 0.$$

The canonical moments of the constrained optimal design are given by $p_{2j-1} = \frac{1}{2}$ $(j = 1,\dots,m+k); p_{2m+2k} = 1$ *and*

$$p_{2l} = \frac{1}{2} \tag{3.15}$$

$l = 1,\dots,m-j-1$

$$p_{2l} = \frac{m+k-l+2}{2(m+k-l+1)} \tag{3.16}$$

$l = m+1,\dots,m+k-1$

$$p_{2l} = \begin{cases} \frac{1}{2}\frac{(2c-1)(l-m+j)-2c}{(2c-1)(l-m+j+1)-2c} & \textit{if } c > \frac{1}{2} \\ \frac{1}{2} & \textit{if } c \leq \frac{1}{2} \end{cases} \tag{3.17}$$

$l = m-j,\dots,m-1$

$$p_{2m} = \begin{cases} \frac{1}{2} + \frac{1}{2}\sqrt{\frac{j+k+1-2(j+k)c}{(k+1)(j+1-2jc)}} & \textit{if } c > \frac{1}{2} \\ \frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{2ck}{k+1}} & \textit{if } c \leq \frac{1}{2}. \end{cases} \tag{3.18}$$

Proof. The equation in (3.15) coincides with (3.5). In the case of equal bounds for the efficiencies $c_l = c; l = m-j,\dots,m-1,m+1,\dots,m+k$, the recursive relation (3.7) simplifies to $p_{2m+2k} = 1$

$$p_{2l} = 1 - \frac{1}{4p_{2l+2}} \qquad l = m+k-1, m+k-2,\dots,m+1 \tag{3.19}$$

which gives (3.16), by induction. If $c \leq \frac{1}{2}$ a further induction and (3.6) show

$$p_{2l} = \frac{1}{2} \qquad l = m-j,\dots,m-1 \tag{3.20}$$

and we obtain from (3.8), (3.16) and (3.20)

$$p_{2m} = \frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{2ck}{k+1}}$$

which coincides with (3.18) for $c \le 1/2$. In this case the constrained optimal design problem is solvable and the canonical moments of the optimal design are specified by (3.15) - (3.18) (for $c \le \frac{1}{2}$).
For $c > \frac{1}{2}$ it follows that (3.15) and (3.16) are still valid by the same reasoning as given in the case $c \le \frac{1}{2}$. For a proof of the equations in (3.17) in the case $c > \frac{1}{2}$ we show by induction that a solution of (3.6) in the interval $(0,1)$ necessarily satisfies

$$(3.21) \qquad p_{2m-2j+2t} = \frac{1}{2}\frac{(2c-1)t-2c}{(2c-1)(t+1)-2c}; \qquad \frac{1}{2} < c < \frac{t+2}{2t+2} \qquad t = 0,\ldots,j-1.$$

Obviously (3.6) reduces to (3.21) for $t=0$ if $c > \frac{1}{2}$. In order to prove the step from $t \to t+1$ we note that

$$(3.22) \qquad \begin{aligned} \prod_{l=m-j}^{m-j+t} p_{2l}q_{2l} &= \frac{1}{4^{t+1}} \prod_{l=0}^{t} \frac{(2c-1)l-2c}{(2c-1)(l+1)-2c} \cdot \frac{(2c-1)(l+2)-2c}{(2c-1)(l+1)-2c} \\ &= \frac{1}{4^{t+1}} \frac{2c((2c-1)(t+2)-2c)}{(2c-1)(t+1)-2c} \end{aligned}$$

and (3.6) yields for $l = m-j+t+1; t \le j-2$

$$(3.23) \qquad p_{2m-2j+2t+2} = \max\left\{\frac{1}{2}\frac{(2c-1)(t+1)-2c}{(2c-1)(t+2)-2c}, \frac{1}{2}\right\} = \frac{1}{2}\frac{(2c-1)(t+1)-2c}{(2c-1)(t+2)-2c}$$

where the last identity follows from $c > \frac{1}{2}$ and $c < \frac{t+2}{2t+2}$ (by the induction hypothesis). Finally, $p_{2m-2l+2t+2} < 1$ yields this inequality for $t+1$, i.e. $c < \frac{t+3}{2t+4}$, which proves (3.21) for all $t = 0,\ldots,j-1$. Note that (3.21) implies

$$(3.24) \qquad c < \frac{j+1}{2j}$$

as a necessary condition for the existence of a solution of the constrained optimal design problem. The remaining part of the proof is obtained by the calculation of p_{2m} using formula (3.8) in Theorem 3.3. Observing (3.16) and (3.22) for $t = j-1$ gives

$$p_{2m} = \frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{k}{k+1}\frac{(2c-1)j-2c}{(2c-1)(j+1)-2c}} = \frac{1}{2} + \frac{1}{2}\sqrt{\frac{j+k+1-2c(j+k)}{(k+1)(j+1-2cj)}}$$

where the argument of the squareroot is nonnegative if and only if

$$c \le \frac{k+j+1}{2(k+j)}$$

[here we use the inequality in (3.21) for $t = j-1$, which implies $j+1 > 2cj$].

□

Our next results describe the solution of the constrained optimal design problem (2.5) with equal efficiencies more explicitly by identifying the support points and weights of the design corresponding to the canonical moments given in Corollary 3.4. The proof involves some more sophisticated results about canonical moments and is therefore deferred to the appendix. Throughout this paper

$$U_k(x) = \frac{\sin((k+1)\arccos x)}{\sin(\arccos x)}$$

denotes the kth Chebyshev polynomial of the second kind [see e.g. Chihara (1978)] and $C_j^{(\lambda)}(x,\nu)$ denotes the jth associated ultraspherical polynomial with parameters λ and ν defined by the recursive relations

$$\begin{aligned} C_{-1}^{(\lambda)}(x;\nu) &= 0, \qquad C_0^{(\lambda)}(x;\nu) = 1, \\ (n+\nu+1)C_{n+1}^{(\lambda)}(x;\nu) &= 2(n+\nu+\lambda)xC_n^{(\lambda)}(x;\nu) - (n+\nu+2\lambda-1)C_{n-1}^{(\lambda)}(x;\nu), \end{aligned} \tag{3.25}$$

$(n \geq 0)$ [see Grosjean (1986) or Lasser (1994)]

Theorem 3.5.
(a) *Let ξ^* denote the solution of the constrained optimal design problem (2.5) with equal efficiencies $c_l = c \in (0,1)$ $\quad l = m-j,\dots,m-1,m+1,\dots,m+k,$ where $k > 0$ and*

$$0 < c \leq \frac{j+k+1}{2(j+k)}.$$

ξ^* is supported at the $m+k+1$ zeros $x_0, x_1, \dots, x_{m+k}$ of the polynomial $(x^2-1)Q_{m+k-1}(x)$, where

$$Q_{m+k-1}(x) = \Big[\Big\{xU_k'(x) - p_{2m}U_{k-1}'(x)\Big\}U_{m-1}(x) - q_{2m}U_k'(x)U_{m-2}(x)\Big] \tag{3.26}$$

if $j = 0$ or $j > 0$ and $c \leq \frac{1}{2}$, and

$$\begin{aligned} Q_{m+k-1}(x) = \Bigg[&\Big(xU_k'(x) - p_{2m}U_{k-1}'(x)\Big)\Big(U_{m-j-1}(x)C_j^{(2)}(x;\nu_j) - U_{m-j-2}(x)C_{j-1}^{(2)}(x;\nu_j)\Big) \\ &- \frac{p_{2m-2}q_{2m}}{q_{2m-2}}U_k'(x)\Big(U_{m-j-1}(x)C_{j-1}^{(2)}(x;\nu_{j-1}) - U_{m-j-2}(x)C_{j-2}^{(2)}(x;\nu_{j-1})\Big)\Bigg] \end{aligned} \tag{3.27}$$

if $j > 0$ and $c > \frac{1}{2}$. Here the parameter ν_j is given by

$$\nu_j = \frac{(j+2) - 2c(j+1)}{2c-1} = \nu_{j-1} - 1 \tag{3.28}$$

and p_{2m-2}, p_{2m} are defined in (3.17) and (3.18), respectively. The weights of the constrained optimal design ξ^* at the support points are obtained by the formula

$$\xi^*(\{x_j\}) = \frac{P_{m+k}(x_j)}{\frac{d}{dx}\{(x^2-1)Q_{m+k-1}(x)\}|_{x=x_j}} \qquad j = 0,\dots,m+k \tag{3.29}$$

where the polynomial $P_{m+k}(x)$ is defined by

$$P_{m+k}(x) = k\Big[\Big\{xU_k(x) - \frac{k+1}{k}q_{2m}U_{k-1}(x)\Big\}U_{m-1}(x) - p_{2m}U_k(x)U_{m-2}(x)\Big] \tag{3.30}$$

if $j=0$ or $j>0$ and $c\le\frac12$, and by

$$P_{m+k}(x) \;=\; \frac{k}{j+1-2cj}\Big[\Big\{xU_k(x)-\frac{k+1}{k}q_{2m}U_{k-1}(x)\Big\}\Big\{U_{m-j-1}(x)U_j(x)-2cU_{m-j-2}(x)U_{j-1}(x)\Big\}$$
$$2q_{2m-2}p_{2m}U_k(x)\Big\{U_{m-j-1}(x)U_{j-1}(x)-2cU_{m-j-2}(x)U_{j-2}(x)\Big\}\Big] \tag{3.31}$$

if $j>0$ and $c>\frac12$.

(b) *Let ξ^* denote the solution of the constrained optimal design problem (2.5) with $k=0$ and equal efficiencies $c_l=c\in(0,1)$, $l=m-j,\dots,m-1$, where*

$$0<c<\frac{j+1}{2j}.$$

ξ^* is supported at the $m+1$ zeros $x_0<x_1<\ldots<x_m$ of the polynomial $(x^2-1)U_{m-1}(x)$ if $j=0$ or $j>0$ and $c\le\frac12$, and at the zeros of the polynomial

$$(x^2-1)\Big[U_{m-j-1}(x)C_j^{(2)}(x;\nu_j)-U_{m-j-2}(x)C_{j-1}^{(2)}(x;\nu_j)\Big]$$

if $j>0$ and $c>\frac12$. The weights of ξ^* at the support points are given by $\xi^*(\{\pm1\})=1/(2m)$ and $\xi^*(\{x_l\})=1/m$ $(l=1,\dots,m-1)$ if $j=0$ or $j>0$ and $c\le\frac12$, and are obtained by the formula (3.29) with $k=0$ where the polynomial $P_m(x)$ is defined by

$$P_m(x)=\frac{1}{2(j+1-2cj)}\Big[U_{m-j-1}(x)\Big(U_{j+1}(x)+\tfrac{2(j+1)c-(j+2)}{j-2(j-1)c}U_{j-1}(x)\Big)-2cU_{m-j-2}(x)U_j(x)\Big]$$

if $j>0$ and $c>\frac12$.

Example 3.6. Consider the case $m=2, j=1, k=1$, which corresponds to the D_1-optimal design problem in a quadratic regression with guaranteed efficiencies for testing the highest coefficients in the cubic and linear regression. By Theorem 3.5(a) a solution of the constrained optimal design problem exists if and only if $c\le 75\%$. In this case the support points of the constrained D_1-D_1-optimal design are the zeros of the polynomial $(x^2-1)Q_2(x)$, where

$$Q_2(x)=\begin{cases}4x^2-2q_4 & \text{if } c\le\frac12\\ \frac{2}{1-c}\{x^2-cq_4\} & \text{if } \frac12<c\le\frac34\end{cases}$$

and

$$p_4=\begin{cases}\frac12(1+\sqrt{1-c}) & \text{if } c\le\frac12\\ \frac12(1+\frac12\sqrt{\frac{3-4c}{1-c}}) & \text{if } \frac12<c\le\frac34\end{cases}$$

Here we used that $U_{-1}(x)=0, U_0(x)=1, U_1(x)=2x, C_1^{(2)}(x,\nu_1)=2(\nu_1+2)x/(\nu_1+1)=x/(1-c)$. Similary, we have

$$P_3(x)=\begin{cases}2x(2x^2-1-q_4) & \text{if } c\le\frac12\\ \frac{2x}{1-c}(x^2-1+p_4c) & \text{if } \frac12<c\le\frac34\end{cases}$$

and Theorem 3.5(a) shows that for $c \le 50\%$ the D_1-D_1-constrained optimal design puts masses

$$\frac{p_4}{2(1+p_4)} \quad \text{and} \quad \frac{1}{2(1+p_4)}$$

at the point ± 1 and $\pm\sqrt{(1-p_4)/2}$ where $p_4 = \frac{1}{2}(1+\sqrt{1-c})$. On the other hand, if $50\% \le c \le 75\%$, the constrained D_1-D_1-optimal design puts masses

$$\frac{cp_4}{2(1-cq_4)} \quad \text{and} \quad \frac{1-c}{2(1-cq_4)}$$

at the points ± 1 and $\pm\sqrt{cq_4}$, respectively, where

$$p_4 = \frac{1}{2}(1+\frac{1}{2}\sqrt{\frac{3-4c}{1-c}}).$$

We finally note that for $c = 75\%$ the constrained D_1-D_1-optimal design has masses 3/10 and 1/5 at the points ± 1 and $\pm\sqrt{3/8}$, respectively.

Remark 3.7. It is worthwhile to mention that Theorem 3.5 also contains the solution of the classical D_1-optimal design problem, which can be seen by looking at the limit $c \to 0$ in Theorem 3.5(a). If $c \to 0$ we have from (3.18) $p_{2m} \to 1$ and the polynomials in Theorem 3.5(a) reduce to

$$\begin{aligned} P_{m+k}(x) &= kU_k(x)[xU_{m-1}(x) - U_{m-2}(x)] \\ Q_{m+k-1}(x) &= [xU_k'(x) - U_{k-1}'(x)]U_{m-1}(x) = kU_k(x)U_{m-1}(x) \end{aligned}$$

where the last equality follows by induction and the recursive relation of the Chebyshev polynomials of the second kind. Transferring Theorem 3.5 to the limiting case $c = 0$ shows that the zeros of the polynomial $(x^2-1)U_k(x)U_{m-1}(x)$ determine the support points of the design ξ^* maximizing (2.5) with $c_l = 0$ ($l = m-j, \dots, m-1, m+1, \dots, m+k$) and that the masses are given in (3.29). Now a straightforward calculation shows that $\xi^*(\{x_j\}) = 0$ whenever $U_k(x_j) = 0$. For the remaining support points $x_0, \dots, x_m$ (satisfying $(x^2-1)U_{m-1}(x) = 0$) we obtain

$$\xi^*(\{x_j\}) = \frac{x_jU_{m-1}(x_j) - U_{m-2}(x_j)}{2x_jU_{m-1}(x_j) + (x_j^2-1)U_{m-1}'(x_j)}$$

which gives

$$\xi^*(\{x_j\}) = \begin{cases} \frac{1}{2m} & \text{if } x_j^2 = 1 \\ \frac{1}{m} & \text{if } U_{m-1}(x_j) = 0 \end{cases} .$$

Here we used the identities

$$U_m(1) = (-1)^m U_m(-1) = m+1$$

and

$$(x^2-1)U_{m-1}'(x) = (m-1)xU_{m-1}(x) - mU_{m-2}(x)$$

which follows from the trigonometric representation $U_{m-1}(x) = \sin(m \arccos x)/\sin(\arccos x)$ for the Chebyshev polynomial of the second kind. By the result of Kiefer and Wolfowitz (1959) the design ξ^* is the D_1-optimal design for the polynomial regression model of degree m.

4 D-D_1-constrained optimal designs

In this section we describe the solution of the constrained optimal design problem (2.6) which maximizes the D-optimality criterion for the model of degree m with guaranteed efficiencies for testing the highest coefficients in the models of degree $m-j, m-j+1, \ldots, m+k$ $(j, k \geq 0$ given). The arguments are essentially the same as in the D_1-D_1-case (but substantially more complicated) and for this reason we only state the main results here. For a detailed discussion and a complete proof we refer to Franke (2000). Our first theorem characterizes the solution of the D-D_1-optimal design problem by a system of nonlinear equations.

Theorem 4.1.

(a) *In the case $m - j \geq 2$ a solution ξ^* of the constrained D-D_1-optimal design problem (2.6) exists if and only if there exists a solution $(p_2, \ldots, p_{2m+2k-2}, p_{2m+2k}) \in [\frac{1}{2}, 1)^{m+k-1} \times \{1\}$ of the system of equations*

$$p_{2l} = \max\Big\{1 - \frac{c_{m+k}}{2^{2(m+k-l)} c_l \prod_{i=l+1}^{m+k-1} p_{2i}(1-p_{2i})}, \frac{1}{2}\Big\}, \tag{4.1}$$

$(l = m+k-1, m+k-2, \ldots, m+1)$

$$p_{2(m-j-1-l)} = \frac{(2l+1)p_{2(m-j-1)} - l}{4l p_{2(m-j-1)} - 2l + 1}, \tag{4.2}$$

$(l = 1, \ldots, m-j-2)$

$$p_{2l} = \max\Big\{\frac{2^{-2(l-1)} c_l}{\prod_{i=1}^{l-1} p_{2i}(1-p_{2i})}, \Big[2 - \frac{2p_{2(m-j-1)} - 1}{1 - p_{2(m-j-1)}} \prod_{i=m-j}^{l-1} \frac{p_{2i}}{1-p_{2i}}\Big]^{-1}\Big\}, \tag{4.3}$$

$(l = m-j, \ldots, m)$,

$$c_{m+k} = 2^{2m+2k-2} \prod_{i=1}^{m+k} q_{2i-2} p_{2i} \tag{4.4}$$

(in the case $k = 0$ the equation (4.4) has to be omitted). Moreover, there exists at most one solution of (4.1) – (4.4) in $[\frac{1}{2}, 1)^{m+k-1} \times \{1\}$, which defines the canonical moments of even order of the constrained D-D_1-optimal design, while all canonical moments of odd order of this design are equal to $\frac{1}{2}$.

(b) *In the case $m - j = 1$ a solution of the D-D_1-constrained optimal design problem exists if and only if there exists a minimal integer $n \in \{1, \ldots, m\}$ such that the system of equations*

$$p_{2l} = \max\Big\{1 - \frac{c_{m+k}}{2^{2(m+k-l)} c_l \prod_{i=l+1}^{m+k-1} p_{2i}(1-p_{2i})}, \frac{1}{2}\Big\}, \tag{4.5}$$

$(l = m+k-1, m+k-2, \ldots, m+1)$

$$c_l = 2^{2l-2} \prod_{i=1}^{l} q_{2i-2} p_{2i} \tag{4.6}$$

$l = 1, \dots, n-1$

$$(4.7) \qquad p_{2l} = \max\Big\{ \frac{2^{-2(l-1)} c_l}{\prod_{i=1}^{l-1} p_{2i}(1-p_{2i})}, \Big[2 - \frac{2p_{2n}-1}{1-p_{2n}} \prod_{i=n+1}^{l-1} \frac{p_{2i}}{1-p_{2i}} \Big]^{-1} \Big\},$$

$(l = n+1, \dots, m,)$

$$(4.8) \qquad c_{m+k} = 2^{2m+2k-2} \prod_{i=1}^{m+k} q_{2i-2} p_{2i}$$

has a solution $(p_2, \dots, p_{2m+2k}) \in [\frac{1}{2}, 1)^{m+k-1} \times \{1\}$ *(in the case* $k = 0$ *the equation (4.8) has to be omitted), such that*

$$(4.9) \qquad c_n \le 2^{2n-2} \prod_{i=1}^{n} q_{2i-2} p_{2i}$$

In the case of existence the solution of (4.5) - (4.8) is unique and gives the canonical moments of even order of the constrained D-D_1-optimal design. The canonical moments of odd order of this design are equal to $\frac{1}{2}$.

Note that Theorem 4.1 formally defines one equation for $p_{2m-2j-2}$ [in case (a)] or p_{2n} [in case (b)]. More precisely, consider the situation for $m - j \ge 2$, where $p_{2m+2k-2}, \dots, p_{2m+2}$ are determined by (4.1). Now (4.2) and (4.3) express $p_2, \dots, p_{2m-2j-4}$, $p_{2m-2j}, \dots, p_{2m}$ in terms of $p_{2m-2j-2}$ and (4.4) reduces to an equation with one unknown variable. A similar reasoning applies in case (b), where we obtain for each $n \in \{1, \dots, m\}$ an equation for p_{2n}.
We will conclude this section by a discussion of the case of equal and maximal bounds in the constraints (2.6), that is $c_l = c \in (0, 1)$, $l = m - j, \dots, m + k$. It can be shown by similar but tedious arguments as given in Section 3 [see Franke (2000)] that a solution of the constrained optimal design problem (2.6) with equal bounds for the efficiencies exists if and only if

$$c \le \frac{j+k+2}{2(j+k+1)}.$$

The case of equality is of particular interest and discussed in the following Theorem.

Theorem 4.2. *The solution* ξ^* *of the constrained optimal design problem (2.6) with equal efficiencies*

$$c_l = \frac{j+k+2}{2(j+k+1)} \qquad l = m-j, \dots, m+k$$

exists and is supported at the $m + k + 1$ *zeros of the polynomial*

$$(4.10) \qquad H_{m+k+1}(x) = (j+k) T_{k+m+1}(x) + U_{m-j-1}(x) T_{j+k+2}(x) - U_{k+m-1}(x)$$

where $T_l(x)$ *and* $U_l(x)$ *denote the Chebyshev polynomial of the first and second kind, respectively. The masses at the support points are given by*

$$(4.11) \qquad \xi^*(\{x_i\}) = \frac{(j+k+1) U_{m+k}(x_i) - U_{j+k}(x_i) U_{m-j-2}(x_i)}{\frac{d}{dx} H_{m+k+1}(x) \mid_{x=x_i}}$$

$(i = 0, \dots, m+k)$.

Example 4.3. Consider a cubic regression, i.e. $m = 3$, and assume that the experimenter is interested in estimating the parameters in this model but wants to have some possibility for checking the polynomials of degree $4, 3, 2$. In this case the criterion (2.6) with $m = 3, k = j = 1$ is appropriate and a solution of the constrained optimal design criterion exists if and only if $c \leq \frac{2}{3}$. We have from Szegö (1975)

$$\begin{aligned} T_5(x) &= 16x^5 - 20x^3 + 5x \\ T_4(x) &= 8x^4 - 8x^2 + 1 \\ U_1(x) &= 2x \\ U_3(x) &= 8x^3 - 4x \end{aligned}$$

which gives for the polynomials in (4.10)

$$H_5(x) = 48x^5 - 64x^3 + 16x.$$

Consequently, the constrained optimal design ξ^* with respect to the criterion (2.6) (with $c_l = 2/3; l = 2, 3, 4$) is supported at the points

$$-1, -\frac{1}{\sqrt{3}}, 0, \frac{1}{\sqrt{3}}, 1$$

with masses

$$\frac{3}{16}, \frac{3}{16}, \frac{1}{4}, \frac{3}{16}, \frac{3}{16}$$

respectively. The D_1-efficiencies of this design are

$$\text{eff}_1^{D_1}(\xi^*) = 50\%; \text{eff}_l^{D_1}(\xi^*) = 66.67\% \quad l = 2, 3, 4$$

while the D-efficiency for estimating the parameters in the cubic model is given by

$$\text{eff}_3^{D}(\xi^*) = 90.75\%.$$

A natural competitor is the D_1-optimal design $\xi_4^{D_1}$ for the polynomial regression of degree 4 which puts masses

$$\frac{1}{8}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{8}$$

at the points

$$-1, -\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, 1,$$

respectively. This design has D_1-efficiencies

$$\text{eff}_4^{D_1}(\xi_4^{D_1}) = 1; \quad \text{eff}_l^{D_1}(\xi_4^{D_1}) = 50\% \qquad l = 1, 2, 3$$

and D-efficiency

$$\text{eff}_3^{D}(\xi_4^{D_1}) = 78.59\%$$

for the cubic model.

Acknowledgements. The authors are grateful to I. Gottschlich who typed most of this paper with considerable technical expertise. The financial support of the Deutsche Forschungsgemeinschaft (SFB 475, Reduction fo complexity in multivariate data structures) is gratefully acknowledged.

5 Appendix: Proof of Theorem 3.5.

We will only give a proof for the more complicated case (a) $k > 0$. The remaining part (b) is treated exactly in the same way and left to the reader. By Corollary 2.2.4 and Theorem 3.4.1 of Dette and Studden (1997) the measure ξ^* defined by the canonical moments of Corollary 3.4 has finite support $\{x_0, \ldots, x_{m+k}\}$ and its Stieltjes transform is given by

$$S(z) = \int_{-1}^{1} \frac{d\xi^*(x)}{z-x} = \sum_{j=0}^{m+k} \frac{\xi^*(\{x_j\})}{z-x_j} = \frac{\underline{P}_{m+k}(z,q)}{(z^2-1)\bar{Q}_{m+k-1}(z,p)}$$

where $\underline{P}_{m+k}(z,q)$ and $\bar{Q}_{m+k-1}(z,p)$ are the support polynomials of the sequences

$$q_1, q_2, \ldots, q_{2m+2k-1}, 0,$$
$$p_1, p_2, \ldots, p_{2m+2k-1}, 1,$$

respectively [and the canonical moments are defined in (3.15) - (3.18)]. This implies that the support points of ξ^* are given by the zeros of the polynomial $(z^2-1)\bar{Q}_{m+k-1}(z,p)$ and that the weights can be obtained as

$$\xi^*(\{x_i\}) = \frac{\underline{P}_{m+k}(x_i,q)}{\frac{d}{dx}\{(z^2-1)\bar{Q}_{m+k-1}(z,p)\}|_{z=x_i}} \qquad i = 0, \ldots, m+k. \tag{5.1}$$

Therefore it remains to show that $\bar{Q}_{m+k-1}(x,p)$ is proportional to the polynomial $Q_{m+k-1}(x)$ defined in (3.26) and (3.27) [corresponding to the case $c \leq \frac{1}{2}$ or $c > \frac{1}{2}$, respectively] and that the right hand side of (5.1) coincides with (3.29), (3.30) and (3.31). For the calculation of the polynomial $\bar{Q}_{m+k-1}(x,p)$ we use Theorem 4.4.2 in Dette and Studden (1997) and obtain

$$\bar{Q}_{m+k-1}(x,p) = G_k(x)H_{m-1}(x) - p_{2m-2}q_{2m}G_{k-1}(x)H_{m-2}(x) \tag{5.2}$$

where $G_k(x)$ and $H_{m-1}(x)$ are the supporting polynomials (with leading coefficient equal to one) of the sequences

$$\frac{1}{2}, p_{2m}, \frac{1}{2}, \ldots, \frac{1}{2}, p_{2m+2k-2}, \frac{1}{2}, 1, \tag{5.3}$$

$$\frac{1}{2}, p_2, \frac{1}{2}, \ldots, \frac{1}{2}, p_{2m-2}, \frac{1}{2}, 1, \tag{5.4}$$

respectively. For the calculation of these polynomials we distinguish two cases:

- **(i) support points in the case $c \leq \frac{1}{2}$ and $j > 0$ or $j = 0$:** In this case we obtain from Corollary 4.3.3 in Dette and Studden (1997)

  $$H_{m-1}(x) = \frac{1}{2^{m-1}}U_{m-1}(x), \quad H_{m-2}(x) = \frac{1}{2^{m-2}}U_{m-2}(x) \tag{5.5}$$

  where $U_{m-1}(x)$ denotes the Chebyshev polynomial of the second kind. Theorem 2.5.1 and Corollary 2.3.6 in the same reference show that $G_k(x)$ is also the supporting polynomial of the sequence

  $$\frac{1}{2}, \tilde{p}_2, \frac{1}{2}, \ldots, \frac{1}{2}, \tilde{p}_{2k}, \frac{1}{2}, 1$$

with $\tilde{p}_{2j} = q_{2m+2k-2j}$ and that $G_k(x)$ can be calculated recursively by $G_0(x) = 1$, $G_1(x) = x$

$$(5.6) \qquad \begin{aligned} G_{j+1}(x) &= xG_{j-1}(x) - \tilde{p}_{2j}\tilde{q}_{2j+2}G_{j-1}(x) \\ &= xG_{j-1}(x) - q_{2m+2k-2j}p_{2m+2k-2j-2}G_{j-1}(x) \end{aligned}$$

$(j = 1, \dots, k-1)$. Observing (3.16) this gives for $j = 1, \dots, k-2$

$$G_{j+1}(x) = xG_{j-1}(x) - \frac{1}{4}\frac{j(j+3)}{(j+1)(j+2)}G_{j-2}(x),$$

which implies

$$G_j(x) = \frac{1}{(j+1)2^j}C_j^{(2)}(x) \qquad j = 0, \dots, k-1$$

where $C_j^{(\lambda)}(x) = C_j^{(\lambda)}(x, 0)$ denotes the j-th ultrasperical polynomial defined in (3.25) and we used the recursive relation for the monic ultraspherical polynomials [see Chihara (1978)]. Observing that $U_n'(x) = 2C_{n-1}^{(2)}(x)$ [see Szegö (1975)] it now follows that

$$(5.7) \qquad G_{k-1}(x) = \frac{1}{k2^k} \cdot U_k'(x); \quad G_k(x) = \frac{1}{k2^k}\{xU_k'(x) - p_{2m}U_{k-1}'(x)\}$$

and (5.2), (5.5) imply for the polynomial $\bar{Q}_{m+k-1}(x, p)$ in the case $j = 0$ or $c \le \frac{1}{2}, j > 0:$

$$(5.8) \quad \bar{Q}_{m+k-1}(x,p) = \frac{1}{k2^{m+k-1}}\Big[\big(xU_k'(x) - p_{2m}U_{k-1}'(x)\big)U_{m-1}(x) - q_{2m}U_k'(x)U_{m-2}(x)\Big]$$

This proves the assertion regarding the support points in the case $j = 0$ or $j > 0$ and $c \le \frac{1}{2}$.

- **(ii) support points in the case** $j > 0; c > \frac{1}{2}$: In this case we have to apply Theorem 4.4.2 in Dette and Studden (1997) twice. More precisely, we have for the polynomial $H_{m-1}(x)$ in (5.2)

$$(5.9) \qquad H_{m-1}(x) = \tilde{G}_j(x)\tilde{H}_{m-j-1}(x) - p_{2m-2j-2}q_{2m-2j}\tilde{G}_{j-1}(x)\tilde{H}_{m-j-2}(x)$$

where $\tilde{H}_{m-j-1}(x)$ and $\tilde{G}_j(x)$ are the supporting polynomials of the sequences

$$(5.10) \qquad \frac{1}{2}, p_2, \frac{1}{2}, \dots, \frac{1}{2}, p_{2m-2j-2}, \frac{1}{2}, 1,$$

$$(5.11) \qquad \frac{1}{2}, p_{2m-2j}, \frac{1}{2}, \dots, \frac{1}{2}, p_{2m-2}, \frac{1}{2}, 1,$$

respectively, and the canonical moments are defined in (3.15) and (3.17) respectively. Now the same arguments as used in the derivation of (5.5) yield

$$(5.12) \qquad \tilde{H}_l(x) = \frac{1}{2^l}U_l(x) \qquad l = m-j-2, m-j-1.$$

By Theorem 2.5.1 and Corollary 2.3.6 in Dette and Studden (1997) we find that $\tilde{G}_j(x)$ can be obtained recursively as $\tilde{G}_0(x) = 1, \tilde{G}_1(x) = x$

$$\begin{aligned} \tilde{G}_{i+1}(x) &= x\tilde{G}_i(x) - p_{2m-2i-2}q_{2m-2i}\tilde{G}_{i-1}(x) \\ &= x\tilde{G}_i(x) - \frac{1}{4}\frac{(2c-1)(j-i+2)-2c}{(2c-1)(j-i+1)-2c}\frac{(2c-1)(j-i-1)-2c}{(2c-1)(j-i)-2c}\tilde{G}_{i-1}(x) \end{aligned}$$

$(i = 1, \ldots, j-1)$. Comparing this recurrence relation with the corresponding recursive relation for the monic version of the associated ultraspherical polynomials defined by (3.25) yields

$$\tilde{G}_i(x) = \frac{(\nu_j + 1)}{2^i(\nu_j + i + 1)} C_i^{(2)}(x, \nu_j); \quad i = 0, \ldots, j; \tag{5.13}$$

where ν_j is defined in (3.28). Combining this result with (5.9) and (5.12) shows

$$\begin{aligned} H_{m-1}(x) &= \frac{(\nu_j + 1)}{2^{m-1}(\nu_j + j + 1)} \Big\{ U_{m-j-1}(x) C_j^{(2)}(x, \nu_j) \\ &- 2(1-c) \frac{(\nu_j + j + 1)}{\nu_j + j} U_{m-j-2}(x) C_{j-1}^{(2)}(x, \nu_j) \Big\} \\ &= \frac{j + 1 - 2jc}{2^{m-1}} \Big\{ U_{m-j-1}(x) C_j^{(2)}(x, \nu_j) - U_{m-j-2}(x) C_{j-1}^{(2)}(x, \nu_j) \Big\} \end{aligned}$$

and a similar argument gives

$$H_{m-2}(x) = \frac{j - 2c(j-1)}{2^{m-2}} \Big\{ U_{m-j-1}(x) C_{j-1}^{(2)}(x, \nu_{j-1}) - U_{m-j-2}(x) C_{j-2}^{(2)}(x, \nu_{j-1}) \Big\}.$$

Finally, formula (5.2), (5.7) and (5.12) yield for the supporting polynomial of the sequence

$$\frac{1}{2}, p_2, \frac{1}{2}, \ldots, \frac{1}{2}, p_{2m+2k-2}, \frac{1}{2}, 1$$

in the case $c > \frac{1}{2}, j > 0$:

$$\begin{aligned} \bar{Q}_{m+k-1}(x, p) = \tfrac{(j+1-2cj)}{k \cdot 2^{k+m-1}} \Big\{ &[x U_k'(x) - p_{2m} U_{k-1}'(x)] \\ &[U_{m-1-j}(x) C_j^{(2)}(x, \nu_j) - U_{m-j-2}(x) C_{j-1}^{(2)}(x, \nu_j)] \\ &- q_{2m} \tfrac{(2c-1)(j-1)-2c}{(2c-1)j-1} U_k'(x) [U_{m-j-1}(x) C_{j-1}^{(2)}(x, \nu_{j-1}) - U_{m-j-2}(x) C_{j-2}^{(2)}(x, \nu_{j-1})] \Big\} \end{aligned} \tag{5.14}$$

which yields the assertion regarding the support points in the case $c > \frac{1}{2}, j > 0$.

In order to complete the proof of Theorem 3.5(a) we have to find the polynomial $\underline{P}_{m+k}(x, q)$ in the numerator of (5.1) which supports the sequence

$$\frac{1}{2}, q_2, \frac{1}{2}, \ldots, \frac{1}{2}, q_{2m+2k-2}, \frac{1}{2}, 0.$$

A similar argument as given in the proof of Theorem 4.4.2 in Dette and Studden (1997) shows that

$$\underline{P}_{m+k}(x, q) = \bar{G}_{k+1}(x) \bar{H}_{m-1}(x) - q_{2m-2} p_{2m} \bar{G}_k(x) \bar{H}_{m-2}(x) \tag{5.15}$$

where $\bar{H}_{m-1}(x)$ and $\bar{G}_{k+1}(x)$ are the supporting polynomials of the sequences

$$\begin{aligned} &\frac{1}{2}, q_2, \frac{1}{2}, \ldots, \frac{1}{2}, q_{2m-2}, \frac{1}{2}, 1; \\ &\frac{1}{2}, q_{2m}, \frac{1}{2}, \ldots, \frac{1}{2}, q_{2m+2k-2}, \frac{1}{2}, 0; \end{aligned} \tag{5.16}$$

respectively. Moreover, these polynomials are also the supporting polynomials of the reversed sequences

$$\frac{1}{2}, p_{2m-2}, \frac{1}{2}, \ldots, \frac{1}{2}, p_2, \frac{1}{2}, 1;$$

$$(5.17) \qquad \frac{1}{2}, q_{2m+2k-2}, \frac{1}{2}, \ldots, \frac{1}{2}, q_{2m}, \frac{1}{2}, 0;$$

[see Studden (1982a)]. Using Corollary 2.3.6 in Dette and Studden (1997) we obtain for the polynomial supporting the sequence (5.17) the recursive relation $\bar{G}_0(x) = 1, \bar{G}_1(x) = x$

$$\begin{aligned} \bar{G}_{i+1}(x) &= x\bar{G}_i(x) - p_{2m+2k-2i+2}q_{2m+2k-2i}\bar{G}_{i-1}(x) \\ &= \begin{cases} x\bar{G}_i(x) - \frac{1}{4}\bar{G}_{i-1}(x) & \text{if } 1 \le i \le k-1 \\ \\ x\bar{G}_k(x) - \frac{k+1}{2k}q_{2m}\bar{G}_{k-1}(x) & \text{if } i = k \end{cases} \end{aligned}$$

which yields $\bar{G}_i(x) = \frac{1}{2^i}U_i(x)(1 \le i \le k)$

$$(5.18) \qquad \bar{G}_{k+1}(x) = \frac{1}{2^k}[xU_k(x) - \frac{k+1}{k}q_{2m}U_{k-1}(x)].$$

For the calculation of the polynomials $\bar{H}_{m-1}(x)$ and $\bar{H}_{m-2}(x)$ we have to distinguish the different cases for $c \in (0,1)$.

- **(iii) weights in the case $c \le \frac{1}{2}$ and $j > 0$ or $j = 0$:** We obtain from Corollary 3.4 and Corollary 4.4.3 in Dette and Studden (1997) for the supporting polynomials of the sequence (5.16)

  $$(5.19) \qquad \bar{H}_l(x) = \frac{1}{2^l}U_l(x) \quad l = m-1, m-2$$

  and (5.15), (5.18), (5.19) give

  $$(5.20) \ \underline{P}_{m+k}(x,q) = \frac{1}{2^{m+k-1}}\Big\{[xU_k(x) - \frac{k+1}{k}q_{2m}U_{k-1}(x)]U_{m-1}(x) - p_{2m}U_k(x)U_{m-2}(x)\Big\}$$

  if $c \le \frac{1}{2}$ or $j = 0$.

- **(iv) weights in the case $c > \frac{1}{2}$ and $j > 0$:** We apply again Theorem 4.4.2 in Dette and Studden (1997) and obtain for the polynomial supporting the sequence in (5.16)

  $$(5.21) \qquad \bar{H}_{m-1}(x) = \hat{G}_{m-j-1}(x)\hat{H}_j(x) - q_{2m-2j-2}p_{2m-2j}\hat{G}_{m-j-2}(x)\hat{H}_{j-1}(x)$$

  where $\hat{G}_{m-j-1}(x)$ and $\hat{H}_j(x)$ are the supporting polynomials of the sequences

  $$\frac{1}{2}, q_2, \frac{1}{2}, \ldots, \frac{1}{2}, q_{2m-2j-2}, \frac{1}{2}, 1;$$
  $$\frac{1}{2}, q_{2m-2j}, \frac{1}{2}, \ldots, \frac{1}{2}, q_{2m-2}, \frac{1}{2}, 1;$$

respectively. Observing (3.15), (3.17) (for $c > \frac{1}{2}$) and Corollary 2.3.6 in Dette and Studden (1997) gives

$$\hat{G}_l(x) = \frac{1}{2^l} U_l(x) \quad l = m-j-2, m-j-1 \tag{5.22}$$

$$\hat{H}_l(x) = \frac{1}{2^l} U_l(x) \quad l = j, j-1 \tag{5.23}$$

and it follows from (5.21) in the case $c > \frac{1}{2}$

$$\bar{H}_{m-1}(x) = \frac{1}{2^{m-1}} [U_{m-j-1}(x)U_j(x) - 2cU_{m-j-2}(x)U_{j-1}(x)]. \tag{5.24}$$

A similar argument yields

$$\bar{H}_{m-2}(x) = \frac{1}{2^{m-2}} [U_{m-j-1}(x)U_{j-1}(x) - 2cU_{m-j-2}(x)U_{j-2}(x)] \tag{5.25}$$

and we have from (5.18), (5.24), (5.25) and (5.15) in the case $c > \frac{1}{2}$ and $j > 0$

$$\begin{aligned} \underline{P}_{m+k}(x,q) = & \frac{1}{2^{m+k-1}} \Big\{ [xU_k(x) - \frac{k+1}{k} q_{2m} U_{k-1}(x)][U_{m-j-1}(x)U_j(x) \\ & - 2cU_{m-j-2}(x)U_{j-1}(x)] \\ & - 2q_{2m-2}p_{2m}U_k(x)[U_{m-j-1}(x)U_{j-1}(x) - 2cU_{m-j-2}(x)U_{j-2}(x)] \Big\}. \end{aligned} \tag{5.26}$$

The assertion of the Theorem now follows from the representation of the weights in (5.1) and the representation of the polynomials $\bar{Q}_{m+k-1}(x,p)$, $\underline{P}_{m+k}(x,q)$ in (5.8), (5.20) (in the case $c \leq \frac{1}{2}$ or $j = 0$) and (5.14), (5.26) (in the case $c > \frac{1}{2}$). $\square$

References

G.E.P. Box, N.R. Draper (1959). A basis for the selection of a response surface design. J. Amer. Statist. Assoc. 54, 622-654.

T.S. Chihara (1978). An Introduction to Orthogonal Polynomials. Gordon and Breach, New York.

M. Clyde, K. Chaloner (1996). The equivalence of constrained and weighted designs in multiple objective design problems. J. Amer. Statist. Assoc. 91, 1236-1244.

R.D. Cook, C. Nachtsheim (1982). Model-robust, linear-optimal design. Technometrics 24, 49-55.

R.D. Cook, W.K. Wong (1994). On the equivalence of constrained and compound optimal designs. J. Amer. Statist. Assoc. 89, 687-692.

H. Dette (1990). A generalization of D- and D_1-optimal designs in polynomial regression. Ann. Statist. 18, 1784-1805.

H. Dette (1995a). Discussion of the paper "Constrained optimization of experimental design" by V. Fedorov and D. Cook. Statistics 26, 153-161.

H. Dette (1995b). Optimal designs for identifying the degree of a polynomial regression. Ann. Statist. 23, 1248-1266.

H. Dette, W.J. Studden (1997). The Theory of Canonical Moments with Applications in Statistics, Probability and Analysis. Wiley, New York.

T. Franke (2000). D- und D_1-optimale Versuchspläne unter Nebenbedingungen und gewichtete Maximin-Versuchspläne bei polynomialer Regression. Dissertation, Ruhr-Universität Bochum (in German).

C.C. Grosjean (1986). The weight functions, generating functions and miscellaneous properties of the sequences of orthogonal polynomials of the second kind associated with the Jacobi and the Gegenbauer polynomials. J. Comp. Appl. Math. 18, 259-307.

P.G. Guest (1958). The spacing of observations in polynomial regression. Ann. Math. Statist. 29, 294-299.

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

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

J.C. Kiefer, J. Wolfowitz (1959). Optimum designs in regression problems. Ann. Math. Statist. 30, 271-294.

E. Läuter (1974). Experimental design in a class of models. Math. Operat. Statist. 5, 673-708.

R. Lasser (1994). Orthogonal polynomials and hypergroups II: The symmetric case. Trans. Amer. Math. Soc. 341, 749-770.

T.S. Lau (1983). Theory of canonical moments and its applications in polynomial regression I, II. Technical Reports 83-23, 83-24. Purdue Univ.

T.S. Lau (1988). D-optimal designs on the unit q-ball. J. Statist. Plann. Inf. 19, 299-315.

C.M.S. Lee (1988a). Constrained optimal designs. J. Statist. Plann. Inf. 18, 377-389.

C.M.S. Lee (1988b). D-optimal designs for polynomial regression, when lower degree parameters are more important. Utilitas Mathematica 34, 53-63.

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

M. Skibinsky (1967). The range of the $(n+1)$th moment for distributions on $[0,1]$. J. Appl. Prob. 4, 543-552.

M. Skibinsky (1986). Principal representations and canonical moment sequences for distributions on an interval. J. Math. Anal. Appl. 120, 95-120.

S. Stigler (1971). Optimal experimental designs for polynomial regression. J. Amer. Statist. Assoc. 66, 311-318.

W.J. Studden (1980). D_s-optimal designs for polynomial regression using continued fractions. Ann. Statist. 8, 1132-1141.

W.J. Studden (1982a). Optimal designs for weighted polynomial regression using canonical moments. Third Purdue Symposium on Decision Theory and Related Topics, vol. 2, S.S. Gupta, J.O. Berger eds., 335-350.

W.J. Studden (1982b). Some robust type D-optimal designs in polynomial regression. J. Amer. Statist. Assoc. 8, 916-921.

W.J. Studden (1989). Note on some ϕ_p-optimal designs for polynomial regression. Ann. Statist. 17, 618-623.

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