© Board of the Foundation of the Scandinavian Journal of Statistics 1996. Published by Blackwell Publishers Ltd., 108 Cowley Road, Oxford OX4 1JF and 238 Main Street, Cambridge, MA 02142, USA. Vol. 23: 195–208, 1996

Optimal Product Designs for Multivariate Regression with Missing Terms

HOLGER DETTE and INGO RÖDER

Technische Universität Dresden

ABSTRACT. A multivariate polynomial regression on the q-cube is considered where not all products of the explanatory variables are present in the model. Using an approach of Lim & Studden (1988) D-optimal product designs for a large class of models of this type are determined in terms of their canonical moments and the efficiencies of these designs are investigated. It is demonstrated that the D-optimal product design provides an efficient solution of an optimal design problem, which is for nearly all cases unsolved.

Key words: D-optimality criterion, multivariate polynomial regression, missing interactions, product designs

1. Introduction

Consider a multiple regression of degree m in $q \geqslant 1$ variables

$$Y(x) = \alpha_0 + \sum_{i=1}^{q} \alpha_i x_i + \sum_{1 \leqslant i_1 \leqslant i_2 \leqslant q} \alpha_{i_1, i_2} x_{i_1} x_{i_2} + \cdots + \sum_{1 \leqslant i_1 \leqslant \cdots \leqslant i_m \leqslant q} \alpha_{i_1, \ldots, i_m} \prod_{j=1}^{m} x_{i_j} \tag{1}$$

where the controlled variable $x = (x_1, \ldots, x_q)^{\mathrm{T}}$ is chosen from the compact design space $[-1, 1]^q$. The regression functions are the $\binom{m+q}{q}$ multiple monomials up to degree m

$$\prod_{j=1}^{q} x_j^{h_j} \quad \text{with} \quad \sum_{j=1}^{q} h_j \leqslant m. \tag{2}$$

An approximate design is a probability measure on the q-cube $[-1, 1]^q$ with finite support $x_1, \ldots, x_l$ and masses $\xi_1, \ldots, \xi_l$. The weight ξ_j gives the relative proportion of the total observations taken at the point x_j $(j = 1, \ldots, l)$. While properties of designs for the multivariate polynomial regression (1) and product type multivariate models have been studied intensively in the literature (see e.g. Kono, 1962; Farell *et al.*, 1967; Lim & Studden, 1988; Rafajlowicz & Myszka, 1988, 1992; Wong, 1994), optimal designs for "incomplete" multivariate models have not been studied so far. This paper considers the problem of designing an experiment for a multivariate polynomial regression (1) where not necessarily all $\binom{m+q}{q}$ monomials of the form (2) appear in the model. To be more precise let for $h_1, \ldots, h_q \in \{0, \ldots, m\}$ with $\Sigma_{j=1}^{q} h_j \leqslant m$

$$\mathscr{I}_{h_1, \ldots, h_q} = \begin{cases} 1 & \text{if } \Pi_{j=1}^{q} x_j^{h_j} \text{ appears in the multivariate polynomial regression} \\ 0 & \text{else} \end{cases}$$

denote $\binom{m+q}{q}$ given numbers with values 0 or 1 and define a linear model by

$$Y(x) = \sum_{\substack{h_1, \ldots, h_q \in \{0, \ldots, m\} \\ \Sigma_{j=1}^{q} h_j \leqslant m}} \mathscr{I}_{h_1, \ldots, h_q} \left(\alpha_{h_1, \ldots, h_q} \prod_{j=1}^{q} x_j^{h_j} \right). \tag{3}$$

The multivariate polynomial regression (3) has

$$N_{q,m,\mathscr{I}} := \sum_{\substack{h_1,\ldots,h_q \in \{0,\ldots,m\} \\ \Sigma_{j=1}^{q} h_j \leqslant m}} \mathscr{I}_{h_1,\ldots,h_q}$$

parameters and the indicator functions $\mathscr{I}_{h_1,\ldots,h_q}$ specify the monomials defined by (2) which appear in the "incomplete" model (3). As a simple example consider the case $q = 3$, $m = 2$ and the polynomial

$$\alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_1 x_2 + \alpha_5 x_1^2 \tag{4}$$

which emerges from (3) by the special choice

$$\begin{aligned} &\mathscr{I}_{0,0,0} = \mathscr{I}_{1,0,0} = \mathscr{I}_{0,1,0} = \mathscr{I}_{0,0,1} = \mathscr{I}_{1,1,0} = \mathscr{I}_{2,0,0} = 1, \\ &\mathscr{I}_{1,0,1} = \mathscr{I}_{0,1,1} = \mathscr{I}_{0,2,0} = \mathscr{I}_{0,0,2} = 0. \end{aligned} \tag{5}$$

The vector of regression functions $f(x)$ in the model (3) consists of the $N_{q,m,\mathscr{I}}$ monomials $\Pi_{j=1}^{q} x_j^{h_j}$ satisfying $\Sigma_{j=1}^{q} h_j \leqslant m$ and $\mathscr{I}_{h_1,\ldots,h_q} = 1$. For a given design ξ on the q-cube $[-1, 1]^q$ the information matrix is defined by

$$\mathbf{M}(\xi) = \int_{[-1,1]^q} f(x) f(x)^{\mathrm{T}} \, d\xi(x). \tag{6}$$

By statistical considerations a good design maximizes an appropriate (concave) function of the information matrix. A design is called D-optimal if ξ maximizes the determinant of the information matrix $\mathbf{M}(\xi)$. D-optimal designs for the "complete" polynomial regression (1) (or equivalently for model (3) with $\mathscr{I}_{h_1,\ldots,h_q} = 1$ for all $h_1, \ldots, h_q$) have been determined numerically by Farell *et al.* (1967) ($m = 3, q = 2$) and Lim & Studden (1988) ($m = 3, 4, 5, q = 2; m = q = 3$). For larger values of m, q the numerical difficulties increase rapidly and the last named authors proposed optimal product designs in order to produce at least efficient designs for the "complete" model (1). In this paper we will deterimine D- and D_s-optimal product designs for many of the "incomplete" models defined in (3). Our main tool is the theory of canonical moments (see Studden, 1980, 1982a, b) which allows an explicit description of the D-optimal product design by the canonical moments of its factors. Some tools for the calculation of the support points and weights of a design from its canonical moments are given in the Appendix in order to make the paper self-contained. Special examples are presented in section 3 while section 4 investigates the multivariate polynomial regression (1) where some of the "highest" terms x_j^m are not present in the model. Efficiency calculations indicate that the observations of Lim & Studden (1988) with respect to the excellent efficiencies of optimal product designs in the "complete" model carry over to the "incomplete" models of the form (3). The results of this paper have been implemented in the program OPTIMAL under MS-DOS which allows an efficient calculation of D-optimal product designs for the model (3) and is available from the authors. For a more detailed description of the program see Röder (1994).

2. D- and D_s-optimal product designs for "incomplete" models

Let $\eta = \xi_1 \times \cdots \times \xi_q$ denote a product design on the q-cube $[-1, 1]^q$ (which means that ξ_j is a probability measure on the interval $[-1, 1]$, $j = 1, \ldots, q$) and define Ξ_q as the set of all product designs on $[-1, 1]^q$. A D-optimal product design for the "incomplete" polynomial regression (3) is a solution of the problem

$$\text{maximize} \quad \det(\mathbf{M}(\eta)) \quad \text{with respect to} \quad \eta \in \Xi_q. \tag{7}$$

A standard argument in design theory shows that the D-optimal product design $\eta^* = \xi_1^* \times \cdots \times \xi_q^*$ has symmetric components ξ_j^* and we can restrict ourselves to designs of this type which we call symmetric product designs.

In what follows we denote by $W_{i(j)}(x_j)$ $(i = 0, \ldots, m)$ the monic orthogonal polynomial of degree i with respect to the symmetric measure ξ_j defined by

$$\int_{-1}^{1} W_{k(j)}(x_j) W_{l(j)}(x_j)\, d\xi_j(x_j) = 0 \quad \text{for } l \neq k \tag{8}$$

(here ξ_j is the jth factor of the product design $\eta = \xi_1 \times \cdots \times \xi_q$). The polynomials $W_{i(j)}(x_j)$ can be defined recursively in terms of the canonical moments of the measure ξ_j (see Lim & Studden, 1988). To this end let ξ denote a probability measure on the interval $[-1, 1]$ with moments $c_i = \int_{-1}^{1} x^i\, d\xi(x)$. Let c_i^+ denote the maximum of the ith moments $\int_{-1}^{1} x^i\, d\mu(x)$ over the set of all probability measures having given moments $c_1, \ldots, c_{i-1}$ and let c_i^- denote the corresponding minimum. The ith canonical moment is defined as

$$p_i = \frac{c_i - c_i^-}{c_i^+ - c_i^-} \quad i = 1, 2, \ldots$$

whenever $c_i^- < c_i^+$ and undefined if $c_i^- = c_i^+$. Note that the canonical moments vary in the interval $[0, 1]$ because $c_i^- \leqslant c_i \leqslant c_i^+$. A design ξ is symmetric if and only if the canonical moments of odd order satisfy $p_{2i-1} = 1/2$ whenever they are defined (see Lau & Studden, 1985). For more details we refer the reader to the work of Studden (1980, 1982a, b). If $p_i^{(j)}$ denotes the ith canonical moment of the symmetric factor ξ_j, then the monic orthogonal polynomials with respect to the measure $d\xi_j(x)$ satisfy the recursion $(q_0^{(j)} = 1)$

$$W_{i+1(j)}(x_j) = x_j W_{i(j)}(x_j) - q_{2i-2}^{(j)} p_{2i}^{(j)} W_{i-1(j)}(x_j) \quad i \geqslant 0 \tag{9}$$

with initial conditions $W_{-1(j)}(x_j) = 0$, $W_{0(j)}(x_j) = 1$ (see Lim & Studden, 1988). Note that $q_i^{(j)}$ is defined by $q_i^{(j)} := 1 - p_i^{(j)}$.

Let $f(x)$ denote the vector of regression functions in the model (3). We define a corresponding vector $g_\eta(x)$ of products of monic orthogonal polynomials as follows. If the lth component of $f(x)$ is $\Pi_{j=1}^q x_j^{h_j}$ $(h_1, \ldots, h_q \in \{0, \ldots, m\}, \Sigma_{j=1}^q h_j \leqslant m, \mathscr{I}_{h_1, \ldots, h_q} = 1)$ then the lth component of g_η is defined as $\Pi_{j=1}^q W_{h_j(j)}(x_j)$ (note that g_η depends on $\eta = \xi_1 \times \cdots \times \xi_q$). In order to determine optimal product designs in the "incomplete" model (3) explicitly we need the following basic assumption.

Assumption 2.1

There exists a symmetric product design η with non-singular information matrix (6) *and there exists a permutation matrix Q such that*

$$Q g_\eta(x) = A_\eta Q f(x) \tag{10}$$

where A_η is a lower triangular matrix with determinant 1.

A necessary and sufficient condition for assumption 2.1 is that, if $\mathscr{I}_{h_1, \ldots, h_q} = 1$, then $\mathscr{I}_{h'_1, \ldots, h'_q} = 1$ for every set $(h'_1, \ldots, h'_q)$ such that $h'_j \leqslant h_j$ and h_j and h'_j have the same parity for $j = 1, \ldots, q$. It also follows readily from the recursive relation (9) that the existence of a symmetric product measure η' (with non-singular information matrix) and a permutation matrix Q satisfying (10) implies that (10) is satisfied for all symmetric product designs $\eta \in \Xi_q$ with $\det M(\eta) > 0$ (with the same matrix Q). As an example consider the model (4) and define $f(x) = (1, x_1, x_2, x_3, x_1 x_2, x_1^2)^{\mathrm{T}}$ then, by the recursion (9), $g_\eta(x) =$

$(1, x_1, x_2, x_3, x_1x_2, x_1^2 - p_2^{(1)})^{\mathrm{T}}$ ($\eta = \xi_1 \times \xi_2 \times \xi_3$) and a suitable choice for Q is $Q = \mathbf{I}_6$ and

$$A_\eta = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ -p_2^{(1)} & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Further illustrations for assumption 2.1 are given in the following section. Throughout this paper we define

$$m_j := \max\{h_j \mid \mathscr{I}_{h_1, \ldots, h_q} = 1\} \tag{11}$$

as the largest exponent of the variable x_j in the incomplete regression (3) ($j = 1, \ldots, q$) and

$$\mathscr{I}(i, j, k) := \sum_{\substack{\Sigma_{l=1}^q h_l \leqslant k \\ h_j = i}} \mathscr{I}_{h_1, \ldots, h_q} \tag{12}$$

as the number of terms in the model (3) which are of degree smaller than or equal to k and which contain the factor x_j^i ($i = 1, \ldots, k, j = 1, \ldots, q$).

Theorem 2.1

The D-optimal product design for an "incomplete" polynomial regression (3) *satisfying assumption 2.1 is given by* $\eta^* = \xi_1^* \times \cdots \times \xi_q^*$. *Here, for* $j = 1, \ldots, q$, ξ_j^* *is a probability measure on the interval* $[-1, 1]$ *which is uniquely determined by its canonical moments*

$$p_{2l-1}^{(j)} = \tfrac{1}{2}, \quad l = 1, \ldots, m_j,$$

$$p_{2l}^{(j)} = \frac{\sum_{i=l}^{m} \mathscr{I}(i, j, m)}{\sum_{i=l}^{m} \mathscr{I}(i, j, m) + \sum_{i=l+1}^{m} \mathscr{I}(i, j, m)}, \quad l = 1, \ldots, m_j - 1, \tag{13}$$

$$p_{2m_j}^{(j)} = 1,$$

where m_j *and* $\mathscr{I}(i, j, m)$ *are defined in* (11) *and* (12).

Proof. The discussion following (7) shows that the optimal product design must be symmetric which implies for the canonical moments of its factors $p_{2i-1}^{(j)} = 1/2$ whenever these are defined. By assumption 2.1 we obtain for the information matrix (6) of a symmetric product design $\eta = \xi_1 \times \cdots \times \xi_q$

$$\mathbf{M}(\eta) = \int_{[-1,1]^d} Q^{\mathrm{T}} A_\eta^{-1} Q g_\eta(x) g_\eta^{\mathrm{T}}(x) Q^{\mathrm{T}} (A_\eta^{\mathrm{T}})^{-1} Q \, d\eta(x)$$

and the determinant of this matrix is given by

$$\begin{aligned} \det(\mathbf{M}(\eta)) &= \det\left(\int_{[-1,1]^d} g_\eta(x)\, g_\eta^{\mathrm{T}}(x)\, d\eta(x)\right) \\ &= \prod_{\Sigma_{l=1}^q h_l \leqslant m} \left(\prod_{j=1}^{q} \left[\int_{-1}^{1} W_{h_j(j)}^2(x_j)\, d\xi_j(x_j)\right]\right)^{\mathscr{I}_{h_1, \ldots, h_q}} \\ &= \prod_{j=1}^{q} \prod_{i=1}^{m} \left[\int_{-1}^{1} W_{i(j)}^2(x_j)\, d\xi_j(x_j)\right]^{\mathscr{I}(i, j, m)} \end{aligned} \tag{14}$$

with

$$\mathscr{I}(i,j,m) = \sum_{\substack{\Sigma_{l=1}^{q} h_l \leq m \\ h_j = i}} \mathscr{I}_{h_1,\ldots,h_q} \quad (i=1,\ldots,m, j=1,\ldots,q).$$

The second line in (14) follows from the first by using the orthogonality relation (8) of the polynomials $W_{i(j)}(x_j)$ which implies that the matrix

$$\int_{[-1,1]^d} g_\eta(x) g_\eta^{\mathrm{T}}(x)\, d\eta(x)$$

is diagonal with elements given by the integrals of the squares of the coefficients of $g_\eta(x)$. The maximization of det $(\mathbf{M}(\eta))$ with respect to the product design $\eta = \xi_1 \times \cdots \times \xi_q$ can now be carried out as a maximization of the factors

$$A_j(\xi_j) = \prod_{i=1}^{m} \left[\int_{-1}^{1} W_{i(j)}^2(x_j)\, d\xi_j(x_j) \right]^{\mathscr{I}(i,j,m)}$$

with respect to the (univariate) measure ξ_j. By lem. 4.3 in Lim & Studden (1988) we obtain

$$A_j(\xi_j) = \prod_{i=1}^{m} \left[2^{2i} \prod_{l=1}^{i} q_{2l-2}^{(j)} p_{2l-1}^{(j)} q_{2l-1}^{(j)} p_{2l}^{(j)} \right]^{\mathscr{I}(i,j,m)} = \prod_{l=1}^{m} [q_{2l-2}^{(j)} p_{2l}^{(j)}]^{\Sigma_{i=l}^{m} \mathscr{I}(i,j,m)}$$

where $p_i^{(j)}$ denotes the ith canonical moments of ξ_i and $p_{2i-1}^{(j)} = 1/2$ (by symmetry of ξ_j). A maximization of $A_j(\xi_j)$ in terms of the even canonical moments $p_{2l}^{(j)}$ yields (13). Note that $\Sigma_{i=l}^{m} \mathscr{I}(i,j,m) = 0$ whenever

$$l > m_j = \max \{h_j \mid \mathscr{I}_{h_1,\ldots,h_q} = 1\}.$$

The uniqueness of the D-optimal product design follows, because $p_{2m_j}^{(j)} = 1$ implies the uniqueness of the single factor ξ_j. □

Theorem 2.1 provides a complete solution of the D-optimal design problem in the class of product designs for all "incomplete" models satisfying assumption 2.1. The optimal design is uniquely characterized in terms of the canonical moments of its factors. The support points and the weights of the jth factor ξ_j with canonical moments (13) can be identified by standard methods (see Studden, 1982a) and a corresponding result is given in the Appendix in order to make the paper self-contained.

In the remaining part of this section we consider an alternative optimality criterion which might be useful if the main interest of the experimenter is the parameters corresponding to the "highest" powers in the model. To this end let $s = m - n$ $(n < m)$ and consider a partition of the information matrix

$$M(\eta) = \begin{pmatrix} M_{11}(\eta) & M_{12}(\eta) \\ M_{21}(\eta) & M_{22}(\eta) \end{pmatrix},$$

where $M_{11}(\eta)$ denotes the information matrix of the design η in the "incomplete" model of degree $n < m$. A D_s-optimal design maximizes

$$|M_{22}(\eta) - M_{21}(\eta) M_{11}^{-1}(\eta) M_{12}(\eta)| = \frac{|M(\eta)|}{|M_{11}(\eta)|}$$

which is proportional to the determinant of the inverse covariance matrix of the least squares estimator for the $s = m - n$ "highest" coefficients $[n < m]$ corresponding to the multiple monomials

$$\sum_{j=1}^{q} x_j^{h_j} \quad \text{with} \quad \mathscr{I}_{h_1,\ldots,h_q} = 1 \quad \text{and} \quad \sum_{j=1}^{q} h_j \in \{n+1, \ldots, m\} \tag{15}$$

in the "incomplete" polynomial regression (3). The following result specifies the canonical moments of the D_s-optimal product design. Here and throughout this paper we use the definition $\Sigma_{i=k}^{l} \beta_i = 0$ whenever $l < k$.

Theorem 2.2

A D_s-optimal product design for the "highest $s = m - n$ ($n < m$) coefficients" specified by (15) in an "incomplete" polynomial regression (3) satisfying assumption 2.1 is given by $\hat{\eta} = \hat{\xi}_1 \times \cdots \times \hat{\xi}_q$. Here, for $j = 1, \ldots, q$, the jth factor $\hat{\xi}_j$ of the D_s-optimal product design $\hat{\eta}$ has canonical moments

$$p_{2l-1}^{(j)} = \tfrac{1}{2}, \quad l = 1, \ldots, m_j,$$

$$p_{2l}^{(j)} = \frac{\sum_{i=l}^{m} \mathscr{I}(i,j,m) - \sum_{i=l}^{n} \mathscr{I}(i,j,n)}{\sum_{i=l}^{m} \mathscr{I}(i,j,m) - \sum_{i=l}^{n} \mathscr{I}(i,j,n) + \sum_{i=l+1}^{m} \mathscr{I}(i,j,m) - \sum_{i=l+1}^{n} \mathscr{I}(i,j,n)}, \tag{16}$$

$$l = 1, \ldots, \min\{n, m_j\},$$

$$p_{2l}^{(j)} = \frac{\sum_{i=l}^{m} \mathscr{I}(i,j,m)}{\sum_{i=l}^{m} \mathscr{I}(i,j,m) + \sum_{i=l+1}^{m} \mathscr{I}(i,j,m)} \quad l = \min\{n, m_j\} + 1, \ldots, m_j,$$

with the convention that this sequence terminates at $p_{2l_0}^{(j)}$ whenever $p_{2l_0}^{(j)}$ is 0 or 1 ($1 \leqslant l_0 \leqslant m_j$) and 0/0 is defined as 0.

Proof. The symmetry (which implies $p_{2l-1}^{(j)} = 1/2$) is obtained by standard arguments. By a similar reasoning as in the proof of theorem 2.1 it follows that the D_s-optimal product design is given by $\hat{\eta} = \hat{\xi}_1 \times \cdots \times \hat{\xi}_q$ where $\hat{\xi}_j$ maximizes the factor

$$B_j(\xi_j) = \frac{\prod_{i=1}^{m} \left(\int_{-1}^{1} W_{i(j)}^2(x_j)\, d\xi_j(x_j) \right)^{\mathscr{I}(i,j,m)}}{\prod_{i=1}^{n} \left(\int_{-1}^{1} W_{i(j)}^2(x_j)\, d\xi_j(x_j) \right)^{\mathscr{I}(i,j,n)}}.$$

Expressing $B_j(\xi_j)$ in terms of canonical moments of the measure ξ_j we obtain by straightforward algebra

$$B_j(\xi_j) = \prod_{l=1}^{n} (q_{2l-2}^{(j)} p_{2l}^{(j)})^{\sigma(l,j,m) - \sigma(l,j,n)} \prod_{l=n+1}^{m} (q_{2l-2}^{(j)} p_{2l}^{(j)})^{\sigma(l,j,m)} \tag{17}$$

with $\sigma(l,j,m) = \Sigma_{i=l}^{m} \mathscr{I}(i,j,m)$. If $n < m_j$, it is easy to see that $\sigma(l,j,m) - \sigma(l,j,n) > 0$ ($l = 1, \ldots, n$) and that $\sigma(l,j,m) > 0$ ($l = n+1, \ldots, m_j$). Thus the assertion follows directly by maximizing (17). The remaining case $n \geqslant m_j$ is more delicate. In this case $B_j(\xi_j)$ reduces to

$$B_j(\xi_j) = \prod_{l=1}^{m_j} (q_{2l-2}^{(j)} p_{2l}^{(j)})^{n_l} \tag{18}$$

where

$$n_l = \sigma(l, j, m) - \sigma(l, j, n) = \sum_{i=l}^{m_j} \mathscr{I}(i, j, m) - \mathscr{I}(i, j, n)$$

$$= \sum_{i=l}^{m_j} \sum_{\substack{n+1 \leqslant \Sigma_{k=1}^q h_k \leqslant m \\ h_j = i}} \mathscr{I}_{h_1, \ldots, h_q}. \tag{19}$$

In the last line we used $\mathscr{I}(i, j, r) = 0$ for $i > m_j$ and $r \in \{n, m\}$ (which is immediate from the definition of m_j and (12)). From (19) we have $n_1 \geqslant n_2 \geqslant \cdots \geqslant n_{m_j} \geqslant 0$ and defining $l_0 + 1 = \min\{l \mid 1 \leqslant l \leqslant m_j, n_j = 0\}$ we obtain from (18)

$$B_j(\xi_j) = \prod_{l=1}^{l_0} (q_{2l-2}^{(j)} p_{2l}^{(j)})^{n_l}.$$

If $l_0 = 0$, then ξ_j can be chosen arbitrarily. We put $p_2^{(j)} = 0$ which corresponds to the assertion of the theorem and obtain the factor design $\hat{\xi}_j$ with mass 1 at the point 0. If $l_0 \geqslant 1$ we obtain $p_{2l_0}^{(j)} = 1$ and $p_{2l}^{(j)} = n_l/(n_l + n_{l+1})$ $(1 \leqslant l \leqslant l_0)$ which is (16) in the case $n \geqslant m_j$. □

Remark 2.1. The proof of theorem 2.2 can quite easily be adapted to the designing problem for the estimation of any subset of parameters in the model (3) which complement a submodel of (3) satisfying assumption 2.1. □

Remark 2.2. Theorems 2.1 and 2.2 consider the design space $[-1, 1]^q$ but the results can easily be transformed to arbitrary cubes of the form $\bigotimes_{j=1}^q [a_j, b_j]$ using a similar argument as in Fedorov (1972, p. 80) and assumption 2.1. The support points of the factors ξ_j have to be transformed linearly while the masses remain unchanged. The canonical moments of the factors ξ_j of the optimal product designs are invariant under such a transformation and the reader can identify the optimal product design directly from (13) and (16) using the results of the Appendix. □

3. Examples

Theorems 2.1 and 2.2 are now illustrated by several examples.

Example 3.1. Consider the model (4). By theorem 2.1 and (5) we obtain for the D-optimal product design $\eta^* = \xi_1^* \times \xi_2^* \times \xi_3^*$ where ξ_2^*, ξ_3^* have canonical moments

$$p_1^{(j)} = \tfrac{1}{2}, \quad p_2^{(j)} = 1 \quad (j = 2, 3)$$

and the canonical moments of ξ_1^* are

$$p_1^{(1)} = p_3^{(1)} = \tfrac{1}{2}, \quad p_2^{(1)} = \tfrac{3}{4}, \quad p_4^{(1)} = 1.$$

The results of theorem A.1 in the Appendix give for the factors of the optimal product design

$$\xi_1^* = \begin{pmatrix} -1 & 0 & 1 \\ \frac{3}{8} & \frac{1}{4} & \frac{3}{8} \end{pmatrix} \quad \xi_2^* = \xi_3^* = \begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

If we are interested in the estimation of the parameters corresponding to the terms $x_1 x_2$ and x_1^2 we apply theorem 2.2 (with $m = 2$, $n = 1$, $m_1 = 2$, $m_2 = 1$, $m_3 = 1$) and obtain for the canonical moments of the factors of the D_1-optimal product design

$$p_1^{(1)} = p_3^{(1)} = \tfrac{1}{2}, \quad p_2^{(1)} = \tfrac{2}{3}, \quad p_4^{(1)} = 1$$

$$p_1^{(2)} = \tfrac{1}{2}, \quad p_2^{(2)} = 1$$

$$p_1^{(3)} = \tfrac{1}{2}, \quad p_2^{(3)} = 0.$$

By theorem A.1 a D_1-optimal product design is given by $\hat{\eta} = \hat{\xi}_1 \times \hat{\xi}_2 \times \hat{\xi}_3$, where

$$\hat{\xi}_1 = \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}, \quad \hat{\xi}_2 = \begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad \hat{\xi}_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Note that this design has a singular information matrix for estimating all parameters in model (4) because the factor $\hat{\xi}_3$ puts mass 1 at the point 0. Other D_1-optimal product designs can be obtained by using arbitrary symmetric measures for the design $\hat{\xi}_3$—see the proof of theorem 2.2. □

Example 3.2. Consider the model $[m = 3, q = 2]$

$$\alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_1^2 + \alpha_4 x_1^3 + \alpha_5 x_1^2 x_2 \tag{20}$$

on the design space $[-1, 1]^2$. Assumption 2.1 is easily verified using the matrix $Q = \mathbf{I}_6$,

$$A_\eta = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ -p_2^{(1)} & 0 & 0 & 1 & 0 & 0 \\ 0 & -p_2^{(1)} - q_2^{(1)} p_4^{(1)} & 0 & 0 & 1 & 0 \\ 0 & 0 & -p_2^{(1)} & 0 & 0 & 1 \end{bmatrix}$$

and the vector of regression functions $f(x) = (1, x_1, x_2, x_1^2, x_1^3, x_1^2 x_2)^{\mathrm{T}}$. For the quantities $\mathscr{I}_{h_1, h_2}$ in the model (3) we have

$$\mathscr{I}_{0,0} = \mathscr{I}_{1,0} = \mathscr{I}_{0,1} = \mathscr{I}_{2,0} = \mathscr{I}_{3,0} = \mathscr{I}_{2,1} = 1$$

$$\mathscr{I}_{0,2} = \mathscr{I}_{1,2} = \mathscr{I}_{0,3} = \mathscr{I}_{1,1} = 0$$

and the D-optimal product design is given by $\eta^* = \xi_1^* \times \xi_2^*$ where the canonical moments of ξ_1^* and ξ_2^* are specified in (13). Thus we obtain

$$p_1^{(1)} = p_3^{(1)} = p_5^{(1)} = \tfrac{1}{2}, \quad p_2^{(1)} = \tfrac{4}{7}, \quad p_4^{(1)} = \tfrac{3}{4}, \quad p_6^{(1)} = 1$$

which gives by theorem A.1 for the first factor of the D-optimal product design

$$\xi_1^* = \begin{pmatrix} -1 & \frac{-1}{\sqrt{7}} & \frac{-1}{\sqrt{7}} & 1 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

Similarly it follows from $p_1^{(2)} = \frac{1}{2}, p_2^{(2)} = 1$ that the second factor is given by

$$\xi_2^* = \begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

In order to calculate the optimal product design for estimating the coefficients of x_1^2, $x_1^2 x_2$ and x_1^3 we apply theorem 2.2 and calculate the canonical moments of the factors of the D_2-optimal product design $[m = 3, n = 1]$ as

$$p_1^{(1)} = p_3^{(1)} = p_5^{(1)} = \tfrac{1}{2}, \quad p_2^{(1)} = \tfrac{1}{2}, \quad p_4^{(1)} = \tfrac{3}{4}, \quad p_6^{(1)} = 1$$

$$p_1^{(2)} = \tfrac{1}{2}, \quad p_2^{(2)} = 1.$$

By theorem A.1 we obtain the D_2-optimal product design for the model (20) as $\hat{\nu} = \hat{\xi}_1 \times \hat{\xi}_2$,

$$\hat{\xi}_1 = \begin{pmatrix} -1 & \frac{-1}{\sqrt{8}} & \frac{-1}{\sqrt{8}} & 1 \\ \frac{3}{14} & \frac{2}{7} & \frac{2}{7} & \frac{3}{14} \end{pmatrix}, \quad \hat{\xi}_2 = \begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}. \qquad \square$$

Example 3.3. Consider the "complete" model (1), i.e. $\mathscr{I}_{h_1,\ldots,h_q} = 1$ for all $h_1, \ldots, h_q \in \{0, \ldots, m\}$ with $\Sigma_{j=1}^q h_j \leq m$. In this case we have

$$\mathscr{I}(i, j, m) = \sum_{\Sigma_{k \neq j} h_k \leq m - i} 1 = \binom{m - i + q - 1}{q - 1}$$

and obtain by theorems 2.1 and 2.2 the results of Lim & Studden (1988). For example, the even canonical moments of the factors of the D-optimal product design are given by

$$p_{2l}^{(j)} = \frac{\sum_{i=l}^{m} \binom{m-i+q-1}{q-1}}{\sum_{i=l}^{m} \binom{m-i+q-1}{q-1} + \sum_{i=l+1}^{m} \binom{m-i+q-1}{q-1}} = \frac{\binom{m-l+q}{q}}{\binom{m-l+q-1}{q} + \binom{m-l+q}{q}}$$
$$= \frac{q+m-l}{q+2(m-l)}, \quad l = 1, \ldots, m$$

independently of $j = 1, \ldots, q$. Here the second line follows from the identity

$$\sum_{j=0}^{i} \binom{q-1+j}{q-1} = \binom{q+i}{q} \tag{21}$$

(see Scheffé, 1958, p. 356).

4. Multivariate regression with missing "highest" terms

Consider the "complete" model, where some of the powers x_j^m are not present in the model. Thus we have in model (3)

$$\begin{aligned}
&\mathscr{I}_{m,0,\ldots,0} = \mathscr{I}_{0,m,0,\ldots,0} = \cdots = \mathscr{I}_{\underbrace{0,\ldots,0}_{k-1},m,0,\ldots,0} = 1 \\
&\mathscr{I}_{0,\ldots,0,m} = \mathscr{I}_{0,\ldots,0,m,0} = \cdots = \mathscr{I}_{\underbrace{0,\ldots,0}_{k},m,0,\ldots,0} = 0 \\
&\mathscr{I}_{h_1,\ldots,h_q} = 1 \quad \text{otherwise}
\end{aligned} \tag{22}$$

and (3) reduces to

$$Y(x) = \alpha_0 + \sum_{j=1}^{q} \alpha_j x_j + \sum_{1 \leq i_1 \leq i_2 \leq q} \alpha_{i_1,i_2} x_{i_1} x_{i_2} + \cdots$$
$$+ \sum_{\substack{1 \leq i \leq \cdots \leq i_m \leq q \\ \exists r \neq s:\, i_r < i_s}} \left(\alpha_{i_1,\ldots,i_m} \sum_{j=1}^{m} x_{i_j} \right) + \sum_{j=1}^{k} \beta_j x_j^m. \tag{23}$$

The quadratic case (i.e. $m = 2$) has been investigated by Uppermann (1993) who found the D-optimal designs explicitly using the Kiefer and Wolfowitz equivalence theorem. Even this case requires an extremely tedious analysis (see Uppermann, 1993) and it seems to be intractable to obtain the optimal designs explicitly for polynomials of higher degree. However, the results of section 2 show that the optimal product designs can be described in terms of the canonical moments of their factors. To this end we observe that assumption 2.1

is satisfied and obtain from (22) and example 4.3

$$\mathscr{I}(i,j,m) = \sum_{\substack{\Sigma_{l=1}^{q} h_l \leqslant m \\ h_j = i}} \mathscr{I}_{h_1,\ldots,h_q} = \binom{m-i+q-1}{q-1} \quad j=1,\ldots,k$$

$$\mathscr{I}(i,j,m) = \begin{cases} \binom{m-i+q-1}{q-1} & \text{if } 1 \leqslant i \leqslant m-1 \\ 0 & \text{if } i=m \end{cases} \quad j=k+1,\ldots,m$$

which implies (using the identity (21))

$$\sum_{i=l}^{m} \mathscr{I}(i,j,m) = \begin{cases} \binom{m-l+q}{q} & \text{if } j=1,\ldots,k \\ \binom{m-l+q}{q} - 1 & \text{if } j=k+1,\ldots,m \end{cases}$$

Now straightforward algebra yields the following result.

Theorem 4.1

The D-optimal product design for the model (23) *is given by*

$$\eta^* = \underbrace{\xi_1^* \times \cdots \times \xi_1^*}_{k} \times \underbrace{\xi_2^* \times \cdots \times \xi_2^*}_{q-k}$$

where the designs ξ_1^* and ξ_2^* are uniquely determined by their canonical moments

$$p_{2l-1}^{(1)} = \tfrac{1}{2}, \quad p_{2l}^{(1)} = \frac{q+m-l}{q+2(m-l)} \quad (l=1,\ldots,m) \quad (p_{2m}^{(1)} = 1)$$

and

$$p_{2l-1}^{(2)} = \tfrac{1}{2} \quad (l=1,\ldots,m-1)$$

$$p_{2l}^{(2)} = \frac{(q+m-l)! - (m-l)!q!}{(q+m-l)! + (m-l)(q+m-l-1)! - 2(m-l)!q!} \quad (l=1,\ldots,m-2)$$

$$p_{2m-2}^{(2)} = 1.$$

The D_s-optimal design can be calculated in a similar way. As an example we consider the case $n = m-1$, which is of particular interest for testing if the degree of the multivariate polynomial (23) is m or $m-1$. The proof of the following result is omitted for the sake of brevity.

Theorem 4.2

The D_1-optimal product design for the "incomplete" multivariate polynomial regression model (23) *is given by*

$$\hat{\eta} = \underbrace{\hat{\xi}_1 \times \cdots \times \hat{\xi}_1}_{k} \times \underbrace{\hat{\xi}_2 \times \cdots \times \hat{\xi}_2}_{q-k}$$

where the designs $\hat{\xi}_1$ and $\hat{\xi}_2$ are uniquely determined by their canonical moments

$$p_{2l-1}^{(1)} = \tfrac{1}{2}, \quad p_{2l}^{(1)} = \frac{q-1+m-l}{q-1+2(m-l)} \quad (l=1,\ldots,m) \quad (p_{2m}^{(1)} = 1)$$

and

$$p^{(2)}_{2l-1} = \tfrac{1}{2} \quad (l = 1, \ldots, m-1)$$

$$p^{(2)}_{2l} = \frac{m-l+q-1-(q-1)!\,\dfrac{(m-l)!}{(m-l-2+q)!}}{2(m-l)+q-1-2(q-1)!\,\dfrac{(m-l)!}{(m-l-2+q)!}} \quad (l = 1, \ldots, m-2)$$

$$p^{(2)}_{2m-2} = 1,$$

respectively.

Example 4.1. Let $m = 3$, $q = 2$, $k = 1$ and consider the model (23) on the square $[-1, 1]^2$, which gives

$$Y(x) = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_1^2 + \alpha_4 x_2^2 + \alpha_5 x_1 x_2 + \alpha_6 x_1 x_2^2 + \alpha_7 x_1^2 x_2 + \alpha_8 x_1^3$$

(compared to the complete model (1) only the term x_2^3 is missing). Theorem 4.1 yields for the D-optimal product design $\eta^* = \xi_1^* \times \xi_2^*$ where the canonical moments of ξ_1^* and ξ_2^* are given by

$$p_1^{(1)} = p_3^{(1)} = p_5^{(1)} = \tfrac{1}{2}, \quad p_2^{(1)} = \tfrac{2}{3}, \quad p_4^{(1)} = \tfrac{3}{4}, \quad p_6^{(1)} = 1$$

$$p_1^{(2)} = p_3^{(2)} = \tfrac{1}{2}, \quad p_2^{(2)} = \tfrac{5}{7}, \quad p_4^{(2)} = 1,$$

respectively. Now theorem A.1 in the Appendix gives for the support points and weights of the designs ξ_1^* and ξ_2^*

$$\xi_1^* = \begin{pmatrix} -1 & \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & 1 \\ \frac{3}{10} & \frac{1}{5} & \frac{1}{5} & \frac{3}{10} \end{pmatrix} \quad \xi_2^* = \begin{pmatrix} -1 & 0 & 1 \\ \frac{5}{14} & \frac{2}{7} & \frac{5}{14} \end{pmatrix}$$

and $\eta^* = \xi_1^* \times \xi_2^*$ is the D-optimal product design. Similarly, it can be shown that the D_1-optimal product design (i.e. the design for estimating the coefficients $x_1 x_2^2, x_1^2 x_2, x_1^3$) is given by $\hat{\eta} = \hat{\xi}_1 \times \hat{\xi}_2$ where

$$\hat{\xi}_1 = \begin{pmatrix} -1 & \frac{-1}{\sqrt{5}} & \frac{-1}{\sqrt{5}} & 1 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix} \quad \hat{\xi}_2 = \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

Note that $\hat{\xi}_1$ and $\hat{\xi}_2$ are the D-optimal designs for a cubic and quadratic univariate polynomial regression (see Atkinson & Donev, 1992, p. 125).

Corollary 4.1

The D_1-optimal product design for the bivariate polynomial regression (23) *with* $q = 2$, $m \in \mathbb{N}$ *and* $k = 1$ *is given by* $\hat{\eta} = \hat{\xi}_1 \times \hat{\xi}_2$ *where* $\hat{\xi}_1$ *and* $\hat{\xi}_2$ *are the D-optimal designs for a univariate polynomial regression of degree* m *and* $m-1$, *respectively.*

Proof. By theorem 4.2 the even canonical moments of $\hat{\xi}_1$ and $\hat{\xi}_2$ are

$$p_{2l}^{(1)} = \frac{m-l+1}{2(m-l)+1} \quad l = 1, \ldots, m$$

and

$$p_{2l}^{(2)} = \frac{m-l}{2(m-l)-1} = \frac{m-1-l+1}{2(m-1-l)+1} \quad l = 1, \ldots, m-1$$

Table 1. *D-efficiencies of the D-optimal product designs η^* for the "incomplete" quadratic model (23) with $q-k$ missing terms. μ^* denotes the D-optimal design*

q	k	$h_{k,q,2}$	det $\mathbf{M}_2(\eta_p^*)$	det $\mathbf{M}_2(\eta^*)$	D-efficiencies
1	1	3	0.148148	0.148148	1
2	1	5	0.105469	0.105469	1
	2	6	0.1112×10^{-1}	0.1143×10^{-1}	0.99553
3	1	8	0.8192×10^{-1}	0.8192×10^{-1}	1
	2	9	0.6711×10^{-2}	0.6815×10^{-2}	0.99830
	3	10	0.5498×10^{-3}	0.5783×10^{-3}	0.99495
4	1	12	0.6698×10^{-1}	0.6698×10^{-1}	1
	2	13	0.4486×10^{-2}	0.4531×10^{-2}	0.99924
	3	14	0.3005×10^{-3}	0.3102×10^{-3}	0.99772
	4	15	0.2013×10^{-4}	0.2157×10^{-4}	0.99539
5	1	17	0.5665×10^{-1}	0.5565×10^{-1}	1
	2	18	0.3210×10^{-2}	0.3232×10^{-2}	0.99962
	3	19	0.1818×10^{-3}	0.1859×10^{-3}	0.99885
	4	20	0.1030×10^{-4}	0.1080×10^{-4}	0.99766
	5	21	0.5840×10^{-6}	0.6348×10^{-6}	0.99604

$(p_{2m}^{(1)} = p_{2m-1}^{(2)} = 1)$. But these are exactly the canonical moments of the D-optimal designs for a (univariate) polynomial regression of degree m and $m-1$ (see Studden, 1980, which proves the assertion. □

It might be of interest to compare the D-optimal product designs of theorem 4.1 with the D-optimal designs calculated by Uppermann (1993) in the quadratic case. To this end we use the D-efficiency

$$\text{eff}\,(\eta^*) = \left(\frac{\det \mathbf{M}(\eta^*)}{\sup_\eta (\det \mathbf{M}(\eta))}\right)^{1/h_{k,q,m}}$$

where η^* is the D-optimal product design, the sup is taken over all designs on the q-cube $[-1, 1]^q$ and $h_{k,q,m}$ is the number of parameters in the model (23) (depending on k, q and m). The efficiencies for various values of the parameters q and k are listed in Table 1. The results indicate that the D-optimal product designs are highly efficient in all considered cases. It is remarkable that in the case $k=1$ the product designs are also globally D-optimal and that the efficiencies of the product designs are slightly decreasing when $h_{k,q,m}$ is increasing.

Acknowledgements

The authors are grateful to three unknown referees for their constructive comments on an earlier version of this paper. We are also indebted to Professor D. Rasch for drawing our attention to the problem of optimal design in "incomplete" multivariate regression models. Parts of the paper were written while the first author was visiting the University of Göttingen and this author would like to thank the Institut für Mathematische Stochastik for its hospitality.

References

Atkinson A. C. & Donev, A. N. (1992). *Optimum experimental designs*, Clarendon Press, Oxford.

Farrell, H. R., Kiefer J., & Walbran, J. (1967). Optimum multivariate designs, *Proc. Fifth Berkeley Symp. Math. Statist. Probab.* 1, 113–138. University of California Press, Berkeley, CA.

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

Kono, K. (1962). Designs for quadratic regression on the k- cube. *Mem. Fac. Sci. Kyushu Univ. Ser. A* **16,** 114–122.

Lau, T.-S. & Studden, W. J. (1985). Optimal designs for trigonometric and polynomial regression using canonical moments. *Ann. Statist.* **13,** 383–393.

Lim, Y. B. & Studden, W. J. (1988). Efficient D_s-optimal designs for multivariate polynomial regression on the q-cube. *Ann. Statist.* **16,** 1225–1240.

Rafajlowicz, E. & Myszka, W. (1988). Optimum experimental design for a regression on a hypercube—generalization of Hoel's result. *Ann. Inst. Statist. Math.* **40,** 821–828.

Rafajlowicz, E. & Myszka, W. (1992). When product type experimental design is optimal? Brief survey and new results. *Metrika* **39,** 321–333.

Röder, I. (1994). Optimale Versuchspläne für Modelle der multivariaten Polynomregression, Diplomarbeit, Technische Universität Dresden.

Scheffé (1958). Experiments with mixtures. *J. Roy. Statist. Soc. Ser. B* **20,** 344–360.

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

Studden, W. J. (1982a). Some robust-type D-optimal designs in polynomial regression. *J. Amer. Statist. Assoc.* **77,** 916–921.

Studden, W. J. (1982b). Optimal designs for weighted polynomial regression using canonical moments. In *Statistical decision theory and related topics III* (eds. S. S. Gupta & J. O. Berger), Vol. 2, 335–350. Academic Press, New York.

Uppermann, P. M. (1993). Designs with a small number of runs for fractorial experiments. Dissertation, Technische Universiteit Eindhoven.

Wong, W. K. (1994). G-optimal designs for multi-factor experiments with heteroscedastic errors. *J. Statist. Planning Inference* **40,** 127–133.

Received November 1994, in final form May 1995

H. Dette, Institut für Mathematische Stochastik, Technische Universität Dresden, Mommsenstrasse 13, 01062 Dresden, Germany.

Appendix

The following result provides an efficient tool for calculating the support points and weights of a design corresponding to a terminating sequence of canonical moments (see Röder, 1994 for a detailed proof).

Theorem A.1

Let ξ denote a probability measure in the interval $[a, b]$, $a, b \in \mathbb{R}$ $(a < b)$ with canonical moments $p_j \in (0, 1)$ for $j = 1, \ldots, 2m-1$ and $p_{2m} = 1$, define $\zeta_1 = p_1$, $\gamma_1 = q_1$, $\zeta_j = q_{j-1}p_j$, $\gamma_j = p_{j-1}q_j$ $(j = 2, \ldots, m)$.

ξ has exactly $m+1$ support points $a = x_0 < x_1 < \cdots < x_{m-1} < x_m = b$, which are the roots of the polynomial

$$\tilde{Q}_{m+1}(x) := (x-a)(x-b)Q_{m-1}(x).$$

The weights at these points are:

$$\xi(\{x_j\}) = \frac{P_m(x_j)}{\frac{d}{dx}\hat{Q}_{m+1}(x)|_{x=x_j}} \quad j = 0, \ldots, m$$

Here the polynomials $P_j(x)$, $Q_j(x)$ are defined recursively by $Q_{-1}(x) = P_{-1}(x) = 0$,

$$Q_0(x) = P_0(x) = 1$$

$$Q_{j+1}(x) = [x - a - (b-a)(\gamma_{2j+2} + \gamma_{2j+3})]Q_j(x) - (b-a)^2\gamma_{2j+1}\gamma_{2j+2}Q_{j-1}(x) \qquad (0 \leqslant j \leqslant m-2)$$

$$P_{j+1}(x) = [x - a - (b-a)(\zeta_{2j+2} + \zeta_{2j+3})]P_j(x) - (b-a)^2\zeta_{2j+1}\zeta_{2j+2}P_{j-1}(x) \qquad (0 \leqslant j \leqslant m-2),$$

and

$$P_m(x) = [x - a - (b-a)\zeta_{2m}]P_{m-1}(x) - (b-a)^2\zeta_{2m-1}\zeta_{2m}P_{m-2}(x).$$