

Optimal designs for dual response polynomial regression models

Fu-Chuen Chang^a, Mong-Na Lo Huang^a, Dennis K.J. Lin^{b, *},
Huic-Ching Yang^a

^a*Department of Applied Mathematics, National Sun Yat-sen University, Taiwan, ROC*

^b*Management Science and Information Systems, The Pennsylvania State University, PA, USA*

Received 18 February 1999; received in revised form 30 May 2000; accepted 16 June 2000

Abstract

In this paper, the D - and D_s -optimal design problems in linear regression models with a one-dimensional control variable and a k -dimensional response variable are considered. The response variables are correlated with a known covariance matrix. Some of the D - and D_s -optimal designs with polynomial models for $k=2$ are found explicitly. It is noted that the number of support points for the D - and D_s -optimal designs highly depend on the correlation between the two response variables except on some special cases. © 2001 Elsevier Science B.V. All rights reserved.

MSC: 62K05

Keywords: Correlated observations; Dual response model; Equivalence theorem; Polynomial regression

1. Background

Shadow is an important component of a cathode-ray tube, which is a main part in a monitor or TV set. The quality of the shadow mask affects the quality of the screen image of its primary product. Since 1984, L-G Micros Co. has produced shadow masks for computer monitors and color TV sets. It has become one of the leading manufacturers of shadow masks. In general, the shadow mask manufacturing process consists of four major steps: coating, exposing, developing and etching. Coating is a process to remove impurities such as oil and dirt from the iron plate (the raw material) and to coat the plate with photosensitive film. Exposing is a process to print the shape of the mask on the coated plate using light and mask patterns. Developing is a process to remove the unexposed area during exposing process, using developing liquids. Etching

* Corresponding author. Tel.: +1-814-865-0377; fax: +1-814-863-2381.

E-mail address: dk15@psu.edu (D.K.J. Lin).

is the final process to make holes on the plate by eroding the exposed area, using eroding liquids. Finally, the iron plate is trimmed into a series of final form of shadow masks.

The major interest is the process yield, namely, the percentage of “good” masks produced. In one of the criteria to determine the goodness of a produced mask, two response variables, the size of the hole (y_1), measured by the radius of circle on the surface of the hole, and the depth of the hole (y_2), were evaluated to see if they meet the target values. It is of interest to find the optimal setting of line speed (the input variable X) which can make the two response variables meet the target ranges. In order to do so, it is important to understand the relationships between the response variables and the input variable, although the function forms describing the regression relationships between them are complicated. The engineers believe that at the current settings x_0 , $f_1(x_0) = f_2(x_0)$ and $f'_1(x_0) = f'_2(x_0)$. That is, equality of the rate of changes for the response at x_0 can be assured. We are interested in finding an optimal experimental design for the company.

2. Preliminaries

After taking the Taylor expansion of the response functions at the point x_0 up to order r and m , respectively, we have the following dual response polynomial models with common intercept and slope as

$$\begin{aligned} E(y_1) &= \eta_1(x, \vartheta) = \theta_0 + \theta_1 x + \theta_{2,1} x^2 + \dots + \theta_{r,1} x^r, \\ E(y_2) &= \eta_2(x, \vartheta) = \theta_0 + \theta_1 x + \theta_{2,2} x^2 + \dots + \theta_{m,2} x^m, \end{aligned} \tag{2.1}$$

where $x \in \Omega = [-1, 1]$, and

$$\text{Cov}(Y) = \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad -1 < \rho < 1.$$

The responses $\eta_i(x, \vartheta)$, $i = 1, 2$, are linear regression functions and can be represented as $\eta_i(x, \vartheta) = f_i^T(x)\vartheta$, with $\vartheta^T = (\theta_0, \theta_1, \theta_{2,1}, \dots, \theta_{r,1}, \theta_{2,2}, \dots, \theta_{m,2})$, $f_1^T(x) = (f_{11}^T(x), f_{12}^T(x), \mathbf{0}_{m-1}^T)$ and $f_2^T(x) = (f_{11}^T(x), \mathbf{0}_{r-1}^T, f_{22}^T(x))$ where $f_{11}^T(x) = (1, x)$, $f_{12}^T(x) = (x^2, \dots, x^r)$, $f_{22}^T(x) = (x^2, \dots, x^m)$ and $\mathbf{0}_j$ is a j -dimensional zero vector.

A design ξ in this case is a probability measure with finite support on Ω . If design ξ has finite support points $x_1, x_2, \dots, x_n \in \Omega$ and concentrates masses $\xi(x_i) = w_i$ at x_i , $i = 1, \dots, n$, where $w_i N = n_i$ are integers, $i = 1, \dots, n$, then the design ξ is called an exact design; if n_i is not restricted to be an integer, then ξ is called an approximate design and will be denoted hereafter as

$$\xi = \left\{ \begin{matrix} x_1 & x_2 & \dots & x_n \\ w_1 & w_2 & \dots & w_n \end{matrix} \right\}.$$

Denote a matrix $\mathbf{F}(x) = [f_1(x), f_2(x)]$. The information matrix for a design ξ with finite support points is

$$\mathbf{M}(\xi) = \int_{\Omega} \mathbf{F}(x) \boldsymbol{\Sigma}^{-1} \mathbf{F}^T(x) d\xi(x),$$

where $\boldsymbol{\Sigma}$ is the common covariance matrix of $\epsilon_j = (\epsilon_{1j}, \dots, \epsilon_{kj})'$ and observations that belong to different experimental runs are assumed to be independent. Let $\hat{\vartheta}$ be the Gauss–Markov estimator of ϑ , then

$$\text{Cov}(\hat{\vartheta}) \propto \mathbf{M}^{-1}(\xi).$$

Without loss of generality, we consider only designs with nonsingular information matrices here (see also Section 4 below).

A design ξ^* is called ϕ_p -optimal if ξ^* maximizes ϕ_p function of the information matrix $\mathbf{M}(\xi)$ among the class of all designs with nonsingular matrices, that is

$$\phi_p(\mathbf{M}(\xi^*)) = \max_{\xi} \phi_p(\mathbf{M}(\xi)), \quad -\infty < p \leq 1,$$

where

$$\phi_p(\mathbf{M}(\xi)) = \begin{cases} \left(\frac{1}{\ell} \text{tr} \mathbf{M}^p(\xi) \right)^{1/p} & (-\infty < p \leq 1, p \neq 0), \\ (\det \mathbf{M}(\xi))^{1/\ell} & (p = 0), \end{cases}$$

where ℓ is the number of parameters in the model.

The equivalence theorem for the ϕ_p -criterion, $-\infty < p \leq 1$ states that a design ξ^* is ϕ_p -optimal if and only if for all $x \in \mathcal{X}$,

$$\text{tr} \mathbf{M}^{-1+p}(\xi^*) \mathbf{F}(x) \boldsymbol{\Sigma}^{-1} \mathbf{F}^T(x) \leq \text{tr} \mathbf{M}^p(\xi^*). \tag{2.2}$$

Moreover, equality in (2.2) is attained at the support points. The most commonly used criterion is probably the one with $p = 0$, known as D -optimality. More details about the theory of optimal designs for linear regression models can be found in Pukelsheim (1993).

In Section 3, some results of D -optimal designs for certain low degree polynomial models in dual response variables are presented, where the polynomial models for the dual responses have special forms to incorporate the special feature appearing in our example. The D_s -optimal design is typically used for testing if certain parameters are equal to zero (see Studden, 1980). This will be discussed in Section 4. In Section 5, a case where the D - and D_s -optimal designs are independent of the covariance matrix is discussed and some useful characteristics of the D -optimal designs with general k response models are presented. Some concluding remarks are given in Section 6.

3. D -optimal designs

A close look at model (2.1) indicates that if both responses exhibit strong linear relationship with the input variable x , then according to the assumptions of our model,

they both are simple linear regression models and are exactly the same. The case with general k responses having exactly the same models will be discussed in Section 5. In this section we will only discuss cases with $(r, m) = (1, 2)$ or $(1, 3)$. The case $(r, m) = (1, m)$ with $m > 1$, for example, represents that the first response function may have shown a strong linear pattern but the second response function may have a curvilinear pattern. Thus, a second- or higher-order polynomial approximation to the true functions of the second response model may be used.

For model (2.1) with $(r, m) = (1, m)$, the information matrix of a design ξ is symmetric and can be written as

$$M(\xi) = \frac{1}{1 - \rho^2} \begin{pmatrix} 2(1 - \rho)M_{11}(\xi) & (1 - \rho)M_{12}(\xi) \\ \cdot & M_{22}(\xi) \end{pmatrix},$$

where

$$\begin{aligned} M_{11}(\xi) &= \int_{-1}^1 f_{11}(x)f_{11}^T(x) d\xi(x) = (c_{i+j-2})_{1 \leq i, j \leq 2}, \\ M_{12}(\xi) &= \int_{-1}^1 f_{11}(x)f_{22}^T(x) d\xi(x) = (c_{i+j})_{1 \leq i \leq 2, 1 \leq j \leq m-1}, \\ M_{22}(\xi) &= \int_{-1}^1 f_{22}(x)f_{22}^T(x) d\xi(x) = (c_{2+i+j})_{1 \leq i, j \leq m-1}, \end{aligned} \tag{3.1}$$

where $c_j = \int_{-1}^1 x^j d\xi(x)$ is the j th moment of the design ξ .

From the equivalence theorem introduced in Section 2, for any design ξ , denote $d_{0,\rho}(\xi, x) = \text{tr}[M^{-1}(\xi)F(x)\Sigma^{-1}F^T(x)]$. Then design $\xi_{0,\rho}^{(r,m)}$ is D -optimal if and only if

$$d_{0,\rho}(\xi_{0,\rho}^{(r,m)}, x) \leq (m + r), \quad \forall x \in \Omega, \tag{3.2}$$

where $(m + r)$ is the number of the unknown parameters, and equality holds at the support points. In the following, the D -optimal designs with $(r, m) = (1, 2)$ or $(1, 3)$ for different values of ρ are given in Theorems 3.1 and 3.2, and the proofs are given in the appendix.

Theorem 3.1. Consider model (2.1) with $(r, m) = (1, 2)$. For a given ρ , $|\rho| < 1$, $\xi_{0,\rho}^{(1,2)}$ is D -optimal if either

- (1) $\rho \geq -\frac{1}{3}$ and $\xi_{0,\rho}^{(1,2)} = \left\{ \begin{matrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{matrix} \right\}$,
- (2) $\rho < -\frac{1}{3}$ and $\xi_{0,\rho}^{(1,2)} = \left\{ \begin{matrix} -1 & 0 & 1 \\ \frac{2}{3(1-\rho)} & \frac{-1-3\rho}{3(1-\rho)} & \frac{2}{3(1-\rho)} \end{matrix} \right\}$.

Theorem 3.2. Consider model (2.1) with $(r, m) = (1, 3)$. For a given ρ , $|\rho| < 1$, $\xi_{0,\rho}^{(1,3)}$ is D -optimal if either

- (1) $\rho \geq -\frac{1}{2}$ and $\xi_{0,\rho}^{(1,3)} = \left\{ \begin{matrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{matrix} \right\}$,

Table 1

Values of the corresponding p_2, p_4, \sqrt{u} and $(1 - \alpha)/2$ for the D -optimal designs $\xi_{0,\rho}^{(1,3)}$ with $\rho < -2/3$

ρ	p_2	p_4	\sqrt{u}	$\frac{1-\alpha}{2}$
-0.70	0.742568	0.963538	0.164546	0.132298
-0.75	0.728790	0.908734	0.257903	0.145267
-0.80	0.711210	0.854497	0.321688	0.161062
-0.85	0.689242	0.801938	0.369476	0.179943
-0.90	0.662791	0.752396	0.405104	0.201706
-0.95	0.632561	0.707082	0.430452	0.225503

$$(2) \quad -\frac{2}{3} \leq \rho < -\frac{1}{2} \text{ and } \xi_{0,\rho}^{(1,3)} = \left\{ \begin{array}{ccc} -1 & 0 & 1 \\ \frac{3}{4(1-\rho)} & -\frac{1+2\rho}{2(1-\rho)} & \frac{3}{4(1-\rho)} \end{array} \right\},$$

$$(3) \quad \rho < -\frac{2}{3} \text{ and } \xi_{0,\rho}^{(1,3)} = \left\{ \begin{array}{cccc} -1 & -\sqrt{u} & \sqrt{u} & 1 \\ \frac{\alpha}{2} & \frac{1-\alpha}{2} & \frac{1-\alpha}{2} & \frac{\alpha}{2} \end{array} \right\},$$

where $u = p_2q_4$ and $\alpha = p_2p_4/(1 - p_2q_4)$, for $q_i = 1 - p_i$ ($i = 2, 4$),

$$p_4 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}, \tag{3.3}$$

where

$$\begin{aligned} a &= 3q_2[1 - \rho + (1 + \rho)p_2], \\ b &= [-4 + (1 - 4\rho - \rho^2)p_2 + (1 + \rho)(5 + \rho)p_2^2], \\ c &= -(1 + \rho)[1 + (2 + \rho)p_2]p_2 \end{aligned} \tag{3.4}$$

and p_2 is a solution of the following equation:

$$3a_1 + 2a_2p_2 + 5a_3p_2^2 + 3a_4p_2^3 = 0, \tag{3.5}$$

where

$$\begin{aligned} a_1 &= p_4^2(2 - p_4 + \rho p_4), \\ a_2 &= p_4[2(1 + \rho) - (1 - \rho)(5 + \rho)p_4 + 2(1 - 3\rho)p_4^2], \\ a_3 &= q_4p_4(q_4 + 2\rho + \rho^2 - 3\rho p_4), \\ a_4 &= (1 + \rho)q_4^2(1 + \rho - 2p_4). \end{aligned}$$

Table 1 provides numerical results for D -optimal designs by solving Eq. (3.5) for $(r, m) = (1, 3)$. It is interesting to find that as $\rho \rightarrow -1$, the corresponding D -optimal design $\xi_{0,\rho}^{(1,3)}$ converges to the D -optimal design for the cubic polynomial model on $[-1, 1]$.

4. D_s -optimal designs

If the main interest is to test whether the expected values of the two responses y_1 and y_2 are the same, in the case of $r = 1$, the problem would lead to test whether all the parameters in $\vartheta_s = (\theta_{r+1,2}, \dots, \theta_{m,2})'$ are equal to 0. It is natural to find designs that minimize the confidence ellipsoid of the estimator $\hat{\vartheta}_s$, that is to minimize the value of the determinant of the corresponding submatrix $D_{22}(\zeta)$ of $\mathbf{D}(\zeta) = \mathbf{M}^{-1}(\zeta)$. The submatrix $D_{22}(\zeta)$ is proportional to the inverse of

$$\mathcal{C}_{22}(\zeta) = M_{22}(\zeta) - M_{21}(\zeta)M_{11}^{-1}(\zeta)M_{12}(\zeta),$$

and the problem is equivalent to maximizing $\det(\mathcal{C}_{22}(\zeta))$. If $\det(M(\zeta)) \neq 0$ and $\det(M_{11}(\zeta)) \neq 0$, then $\det(\mathcal{C}_{22}(\zeta)) = \det(M(\zeta)) / \det(M_{11}(\zeta))$ (see Graybill, 1983, Theorem 8.2.1). The design ζ_{D_s} is called a D_s -optimal design if

$$\det(\mathcal{C}_{22}(\zeta_{D_s})) = \max_{\zeta} \det(\mathcal{C}_{22}(\zeta)) \tag{4.1}$$

among all designs defined on $[-1, 1]$.

Now for model (2.1), the design $\zeta_{D_s, \rho}^{(r,m)}$ is D_s -optimal if and only if

$$d_{s, \rho}(\zeta_{D_s, \rho}^{(r,m)}, x) = d_{0, \rho}(\zeta_{D_s, \rho}^{(r,m)}, x) - \frac{2}{1 + \rho} f_{11}^T(x) M_{11}^{-1}(\zeta_{D_s, \rho}^{(r,m)}) f_{11}(x) \leq m - r \tag{4.2}$$

for all $x \in [-1, 1]$. Moreover, the equality holds at the support points.

In (4.2), it is assumed that $\det(\mathbf{M}(\zeta_{D_s, \rho}^{(r,m)})) \neq 0$ and $\det(M_{11}(\zeta_{D_s, \rho}^{(r,m)})) \neq 0$, but in some cases, D_s -optimal designs may be degenerate. In those cases, generalized inverses of $\mathbf{M}(\zeta_{D_s, \rho}^{(r,m)})$ and $M_{11}(\zeta_{D_s, \rho}^{(r,m)})$ are considered. Here a matrix B^- is said to be a generalized inverse of a matrix B , if $BB^-B = B$. Moreover if G is any solution to the system $B^2G = B$, the matrix $G^T B G$ is a generalized inverse of B (see Graybill, 1983, Corollary 6.5.2). These are used in some cases in verifying the D_s -optimality of some designs. The following theorems provide D_s -optimal designs for the cases $(r, m) = (1, 2)$ or $(1, 3)$. Their proofs can be obtained by following similar arguments to the proofs given in the appendix for D -optimality and thus are omitted.

Theorem 4.1. Consider model (2.1) with $r = 1$ and $m = 2$. For a given ρ , $|\rho| < 1$, $\zeta_{D_s, \rho}^{(1,2)}$ is D_s -optimal if either

- (1) $\rho \geq 0$ and $\zeta_{D_s, \rho}^{(1,2)} = \left\{ \begin{matrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{matrix} \right\}$, or
- (2) $\rho < 0$ and $\zeta_{D_s, \rho}^{(1,2)} = \left\{ \begin{matrix} -1 & 0 & 1 \\ \frac{1}{2(1-\rho)} & \frac{-\rho}{1-\rho} & \frac{1}{2(1-\rho)} \end{matrix} \right\}$.

Theorem 4.2. Consider model (2.1) with $r = 1$ and $m = 3$. For a given ρ , $|\rho| < 1$, $\zeta_{D_s, \rho}^{(1,3)}$ is D_s -optimal if either

- (1) $\rho \geq -\frac{1}{3}$ and $\zeta_{D_s, \rho}^{(1,3)} = \left\{ \begin{matrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{matrix} \right\}$, or

Table 2

Values of the corresponding p_2, p_4, \sqrt{u} and $(1 - \alpha)/2$ for the D_s -optimal designs $\xi_{D_s, \rho}^{(1,3)}$ with $\rho < -3/5$

ρ	p_2	p_4	\sqrt{u}	$\frac{1-\alpha}{2}$
-0.65	0.79778	0.947051	0.205527	0.105570
-0.70	0.758162	0.896037	0.280750	0.131265
-0.75	0.715092	0.847915	0.329780	0.159837
-0.80	0.66991	0.803546	0.362776	0.190058
-0.85	0.624354	0.763362	0.384377	0.220384
-0.90	0.580118	0.727383	0.397681	0.249380
-0.95	0.53844	0.695304	0.405044	0.276072

$$(2) \quad -\frac{3}{5} \leq \rho < -\frac{1}{3} \text{ and } \xi_{D_s, \rho}^{(1,3)} = \left\{ \begin{array}{ccc} -1 & 0 & 1 \\ \frac{2}{3(1-\rho)} & -\frac{1+3\rho}{3(1-\rho)} & \frac{2}{3(1-\rho)} \end{array} \right\}, \text{ or}$$

$$(3) \quad \rho < -\frac{3}{5} \text{ and } \xi_{D_s, \rho}^{(1,3)} = \left\{ \begin{array}{cccc} -1 & -\sqrt{u} & \sqrt{u} & 1 \\ \frac{\alpha}{2} & \frac{1-\alpha}{2} & \frac{1-\alpha}{2} & \frac{\alpha}{2} \end{array} \right\},$$

where $u = p_2q_4$ and $\alpha = p_2p_4/(1 - p_2q_4)$ for $q_i = 1 - p_i$ ($i = 2, 4$), p_4 is the same as defined in (3.3) and p_2 is a solution of the following equation:

$$4a_1 + 3a_2p_2 + 8a_3p_2^2 + 5a_4p_2^3 = 0 \tag{4.3}$$

with the restriction that $0 < p_2 < 1$, and $a_i, i = 1, \dots, 4$, are as defined in Theorem 3.2.

Some numerical results for D_s -optimal designs obtained by solving Eq. (4.3) for different $\rho, \rho < -3/5$, are listed in Table 2. Similarly as in Section 3, the limit of the D_s -optimal designs should converge to the usual D_s -optimal design on $[-1, 1]$ for estimating the last two parameters.

5. Some extensions

In this section we first discuss the special case when regression functions for both responses are of same degree, i.e. $r = m$ in (2.1). This model arises when the degrees of the response functions are uncertain but do exhibit similar curvilinear patterns. We then approximate both of them by polynomials of the same degree through Taylor expansion. It is interesting to note that in this case, unlike the results give in Section 3, all D - and D_s -optimal designs are independent of the value of ρ , and are given below.

Then for model (2.1) with $r = m \geq 2$, the information matrix of a design ξ is symmetric and can be written as

$$M(\xi) = \frac{1}{1 - \rho^2} \begin{pmatrix} 2(1 - \rho)M_{11}(\xi) & (1 - \rho)M_{12}(\xi) & (1 - \rho)M_{12}(\xi) \\ (1 - \rho)M_{12}^T(\xi) & M_{22}(\xi) & -\rho M_{22}(\xi) \\ (1 - \rho)M_{12}^T(\xi) & -\rho M_{22}(\xi) & M_{22}(\xi) \end{pmatrix},$$

where $M_{ij}(\xi), 1 \leq i \leq j \leq 2$ is as defined in (3.1).

Theorem 5.1. For model (2.1) with $r = m$, the D - and D_s -optimal designs are independent of ρ .

Corollary 5.2. For model (2.1) with $r = m = 2$, then $\xi_{0,\rho}^{(2,2)}$ and $\xi_{D_s,\rho}^{(2,2)}$ are D - and D_s -optimal designs, respectively, where for all $|\rho| < 1$,

$$(1) \xi_{0,\rho}^{(2,2)} = \left\{ \begin{array}{ccc} -1 & 0 & 1 \\ \frac{3}{8} & \frac{1}{4} & \frac{3}{8} \end{array} \right\},$$

$$(2) \xi_{D_s,\rho}^{(2,2)} \text{ is } D_s\text{-optimal with } \xi_{D_s,\rho}^{(2,2)} = \left\{ \begin{array}{ccc} -1 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \right\}.$$

Now, let $d = (d_1, \dots, d_k)$ denote the vector of degrees in the k response model. Next, we present an important invariance property of D -optimal designs for general k response polynomial regression which holds for $k = 1$.

Corollary 5.3. For the k response model with $d = (d_1, \dots, d_k)$, the D -optimal design is scale invariant, i.e. the D -optimal designs on the scaled design space $[ab, b]$ can be calculated from the optimal designs on the interval $[a, 1]$ by multiplying the support points by b .

Proof. Let $\xi(x)$ be a design on $[a, 1]$ with support x_1, \dots, x_n and $\xi_{a,b}(bx) = \xi(x)$ a design on $[ab, b]$. Note that the information matrix of $\xi_{a,b}(x)$ has the form

$$M(\xi_{a,b}(x)) = \sum_{i=1}^n \xi(x_i) \mathbf{F}(bx_i) \Sigma^{-1} \mathbf{F}^T(bx_i) = \mathbf{D} M(\xi(x)) \mathbf{D},$$

where $\mathbf{D} = \text{diag}(1, b, \dots, b^{d_1}, b^2, \dots, b^{d_2}, \dots, b^2, \dots, b^{d_k})$. The desired property is established by the fact $\det M(\xi_{a,b}(x)) = (\det \mathbf{D})^2 \cdot \det M(\xi(x))$.

Note that for the multiresponse polynomial regression model $d = (d_1, \dots, d_k)$, $k > 1$, the D -optimal design is not translation invariant. Furthermore, we offer the following results.

Corollary 5.4. Suppose that all the k response functions in a k response model with covariance matrix Σ are the same, i.e. $\mathbf{F}(x) = (f(x), f(x), \dots, f(x))$. Then the design ξ^* is simultaneously D -optimal for the models $f(x)$ and $\mathbf{F}(x)$.

Proof. Let $\Sigma^{-1} = (\sigma_{ij})$. Then the information matrix is of a form

$$M(\xi^*) = \sum \xi^*(x_i) \mathbf{F}(x_i) \Sigma^{-1} \mathbf{F}^T(x_i) = \left(\sum_{ij} \sigma_{ij} \right) \cdot \sum \xi^*(x_i) f(x_i) f^T(x_i)$$

which is proportional to that for the single response model $f(x)$. Thus the desired property is proved. \square

In particular, the D -optimal design ξ^* for the polynomial regression of degree m on the interval $[-1, 1]$ has equal masses at the $m+1$ zeros of the polynomial $(x^2 - 1)P'_m(x)$

where P'_m denotes the derivative of the m th Legendre polynomial (see Fedorov, 1972, Theorem 2.3.3). The D -optimal design $\xi_{a,b}^*$ on the interval $[a, b]$ can be obtained from ξ^* by the linear transformation $\xi_{a,b}^* = \xi^*((2x - b - a)/(b - a))$.

Note that the model presented in Corollary 5.4 can also be thought of as the random block-effects model as discussed in Cheng (1995) with general block size k , but the design points are restricted to be the same for all k response variables. Then by Corollary 5.4 the D -optimal design problem for the random block-effects model with general block size k can be reduced to be a univariate D -optimal design problem.

6. Discussions

Krafft and Schaefer (1992) studied multiresponse linear regression models, where the parameters in each response were different. They proved that multiresponse D -optimal designs under their models do not depend on the covariance matrix of the response variables. A related problem was also discussed in Bischoff (1995). Recently, Cheng (1995) and Atkins and Cheng (1997) derived the exact D - and A -optimal quadratic regression designs under random block-effects models for any block size and block of size 2, respectively.

In this work, we have found the D - and D_s -optimal designs for dual response polynomial models with some common parameters and known covariance structure. The optimal design problems for this type of multiresponse model have not been considered before, and it is interesting to observe that the D - and D_s -optimal designs do depend on the correlation of the two responses except for some special cases. This differs from results by Krafft and Schaefer (1992) and Bischoff (1995); under their models the optimal designs are independent of the covariance structure.

Note that for $\rho \geq 0$, all D - and D_s -optimal designs obtained in Sections 3 and 4 are the same – both are supported equally on the two boundary points only. This reminds us of the work of Abt et al. (1997, 1998) which discussed optimal design problems for parameter estimation and growth prediction under linear and quadratic growth curve model with intraclass and autoregressive correlation structures. The results found there indicated that in many situations there is essentially no difference between the independent and the correlated error structures. In other words, from Abt et al. (1997), the optimal design for the linear growth model is quite robust against correlated error structure when $\rho > 0$.

In this paper, we used elementary analysis based on the equivalence theorem and some of the basic properties of the canonical moments to find the optimal designs for polynomial models of degree three or less. More investigations are needed for finding general patterns for the D - and D_s -optimal designs in higher degree polynomial models. It is clear that the optimal designs would become more complicated as the polynomial degree increases because there would be more critical regions corresponding to the correlation coefficient ρ and the number of support points for the optimal designs would change accordingly. Also, a natural extension is to investigate optimal designs for the general $\phi_p, (-\infty < p < 1)$ criteria.

As discussed in Dette (1997a,b), the general ϕ_p criteria, $-\infty < p < 1$, ($p \neq 0$) depend on the variances of the least-squares estimator for the individual parameters. These variances are of different scale in the polynomial models. Therefore, averaging might not be an appropriate operation in finding optimal designs. Also, except for $p=0$, the optimal designs found under the general ϕ_p -criteria are not invariant under affine transformations. Dette (1997a,b) proposed to use some standardized criteria to fix this problem. With our model, we will need to find optimal designs for estimating each individual coefficient first. This would provide suitable standardization constants in obtaining a properly standardized A -optimal design. This approach will be studied in future research.

Another possible extension would be to consider models with more than two responses. We believe that the optimal designs will usually be heavily dependent on the covariance structure of the responses. Theorem 5.1 provides an exception where the D - and D_s -optimal designs are independent of ρ values in the covariance matrix. This case can be easily extended to $k \geq 3$ response polynomial models of the same degree, but with the first r parameter coefficients being the same and the rest unrestricted. Although a general solution may be difficult to find, we hope that the results given in this paper will be helpful in studying these general cases.

Acknowledgements

We wish to thank the Co-ordinating Editor, an Associate Editor, and two referees for their insightful comments and suggestions that resulted in a much improved paper. This work was done while Dennis Lin visited The Institute of Statistical Sciences, Academia Sinica, under the support of *National Science Council of ROC* via Contract NSC 87-2119-M-001-007. Mong-Na Lo Huang was supported in part by National Science Council of ROC via Contract NSC-88-2118-M-110-004.

Appendix

For any design ξ , by the strictly concave property of $\log \det(M(\xi))$ in ξ (see Fedorov, 1972, Section 5.1), it is easily seen we can restrict our attention to symmetric designs.

Moreover, in model (2.1) for D -optimality, the corresponding dispersion function $d(\xi, x)$ is a polynomial of degree $2m$ for any design ξ . Thus, by the equivalence theorem (2.2), the D -optimal design has at most $m + 1$ support points. Moreover, if the optimal design has $m + 1$ support points, it must contain the two boundary points $\{-1, 1\}$. The result also holds for D_s -optimality. Now we are ready to prove the main theorems.

Proof of Theorem 3.1. In model (2.1) with $r = 1$ and $m = 2$, the determinant of the information matrix of a symmetric design ξ_s can be expressed as

$$\det(M(\xi_s)) = \frac{4}{(1 + \rho)^3(1 - \rho)} c_2 \left(c_4 - \frac{1 - \rho}{2} c_2^2 \right).$$

In order to maximize $\det(\mathbf{M}(\xi_s))$, it is easy to see that we need $c_4 = c_2$. Furthermore, $\det(\mathbf{M}(\xi_s))_{c_4=c_2} = 4(c_2^2 - [(1 - \rho)/2]c_2^3)/[(1 + \rho)^3(1 - \rho)]$ attains its maximum when $c_2 = 4/3(1 - \rho)$. But when $\rho \geq -\frac{1}{3}$, the value $4/3(1 - \rho) \geq 1$; thus, it takes $c_2 = 1$ to maximize $\det(\mathbf{M}(\xi_s))_{c_4=c_2}$. Therefore, if $\rho \geq -\frac{1}{3}$, $\det(\mathbf{M}(\xi_s))$ attains its maximum when $c_2 = c_4 = 1$. Thus the D -optimal design $\xi_{0,\rho}^{(1,2)}$ must be supported at -1 and 1 . Since $\xi_{0,\rho}^{(1,2)}$ is symmetric, when $\rho \geq -\frac{1}{3}$,

$$\xi_{0,\rho}^{(1,2)} = \left\{ \begin{array}{cc} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}.$$

If $\rho < -\frac{1}{3}$, $\det(\mathbf{M}(\xi_s))$ attains its maximum when $c_2 = c_4 = 4/3(1 - \rho)$. Thus the D -optimal $\xi_{0,\rho}^{(1,2)}$ must be supported at $-1, 0, 1$ and $\xi_{0,\rho}^{(1,2)}(1) = 2/3(1 - \rho)$. Then when $\rho < -\frac{1}{3}$,

$$\xi_{0,\rho}^{(1,2)} = \left\{ \begin{array}{ccc} -1 & 0 & 1 \\ \frac{2}{3(1-\rho)} & \frac{-1-3\rho}{3(1-\rho)} & \frac{2}{3(1-\rho)} \end{array} \right\}.$$

Proof of Theorem 3.2. From the equivalence theorem, it is clear the D -optimal design can only be supported by 2, 3, or 4 points:

1. Suppose the D -optimal design is a two-point design and we consider a symmetric design $\xi_s(u)$:

$$\xi_s(u) = \left\{ \begin{array}{cc} -u & u \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}, \quad 0 < u \leq 1.$$

The determinant of the information matrix of $\xi_s(u)$ is

$$\begin{aligned} \det(\mathbf{M}(\xi_s(u))) &= \frac{u^{12}}{(1 - \rho)^2(1 + \rho)^2} \\ &\leq \frac{1}{(1 - \rho)^2(1 + \rho)^2} = \det(\mathbf{M}(\xi_s(1))). \end{aligned}$$

Thus, the design $\xi_s(1)$ maximizes $\det(\mathbf{M}(\xi))$ among all two-point designs on $[-1, 1]$.

By (3.2), $\xi_s(1)$ is D -optimal if and only if

$$g_\rho(x) = d_{0,\rho}(\xi_s(1), x) = \frac{2(1 + x^2)[1 - (1 - \rho)x^2 + x^4]}{1 + \rho} \leq 4 \tag{A.1}$$

for all $x \in [-1, 1]$. It can be easily verified using (A.1) that the design $\xi_s(1)$ is D -optimal when $\rho \geq -\frac{1}{2}$. When $\rho < -\frac{1}{2}$, the D -optimal design is not supported by two points.

2. When $\rho < -\frac{1}{2}$, suppose the D -optimal design is a three-point design. Thus, consider a three point symmetric design

$$\xi_s(u, \alpha) = \left\{ \begin{array}{ccc} -u & 0 & u \\ \frac{\alpha}{2} & 1 - \alpha & \frac{\alpha}{2} \end{array} \right\},$$

where $0 < u \leq 1$ and $0 < \alpha < 1$. The determinant of the information matrix of $\xi_s(u, \alpha)$ is

$$\det(\mathbf{M}(\xi_s(u, \alpha))) = \frac{\alpha^3(2 - \alpha + \rho\alpha)}{(1 - \rho)^2(1 + \rho)^3} u^{12}.$$

To maximize $\det(\mathbf{M}(\xi_s(u, \alpha)))$, it is easy to see that we need $u = 1$. Moreover, $\det(\mathbf{M}(\xi_s(1, \alpha))) = [1/(1 - \rho)^2(1 + \rho)^3]\alpha^3(2 - \alpha + \rho\alpha)$ attains its maximum at $\alpha_\rho^* = 3/2(1 - \rho)$. Hence, when $\rho < -\frac{1}{2}$, the design $\xi_s(1, \alpha_\rho^*)$ maximizes $\det(\mathbf{M}(\xi))$ among all three-point designs on $[-1, 1]$.

By (3.2), the design $\xi_s(1, \alpha_\rho^*)$ is D -optimal if and only if

$$\begin{aligned} h_\rho(x) &= d_{0,\rho}(\xi(1, \alpha_\rho^*), x) \\ &= \frac{4[3(1 + \rho) - (1 - \rho)(2 + 3\rho)x^2 + (1 - \rho)(1 + 3\rho)x^4 + (1 - \rho)x^6]}{3(1 + \rho)} \\ &\leq 4 \end{aligned} \tag{A.2}$$

for all $x \in [-1, 1]$. Thus, when $\rho \geq -\frac{2}{3}$, it can be easily checked that $h_\rho(x)$ has a unique local maximum on $[-1, 1]$ and $\xi_s(1, \alpha_\rho^*)$ is indeed D -optimal when $-\frac{2}{3} \leq \rho < -\frac{1}{2}$. When $\rho < -\frac{2}{3}$, $h_\rho(x)$ has a local minimum at $x = 0$, so $\xi_s(1, \alpha_\rho^*)$ is not D -optimal.

- When $\rho < -\frac{2}{3}$, the D -optimal design is a four-point design. The determinant of the information matrix of a four-point symmetric design ξ_s can be expressed as

$$\det(\mathbf{M}(\xi_s)) = \frac{[2c_4 - (1 - \rho)c_2^2][2c_2c_6 - (1 - \rho)c_4^2]}{(1 - \rho)^2(1 + \rho)^4}. \tag{A.3}$$

In Studden (1980), canonical moments corresponding to a design have been defined and used to find D_s -optimal designs for polynomial regression. Since it is a very useful tool which can take away the restrictions among the c_i 's, we use it here to simplify the computation of the D -optimal design. Now, let c_i^+ and c_i^- denote the maximum and minimum of the i th moment $\int_{-1}^1 x^i d\xi(x)$ among the set of designs on $[-1, 1]$ whose moments up to the order $i - 1$ coincide with the given moments $\mathbf{c}_{i-1} = (c_1, \dots, c_{i-1})$. The canonical moments are defined by $p_i = (c_i - c_i^-)/(c_i^+ - c_i^-)$, $i = 1, 2, \dots$. Then for a symmetric design ξ_s with four points defined on $[-1, 1]$, from Studden (1980), it can be seen that $c_{2i-1} = 0$, $p_{2i-1} = 1/2$, $i = 1, \dots, 3$, as well as correspondence between c_{2i} and p_{2i} , $i = 1, \dots, 3$ with $0 \leq p_i \leq 1$, and $q_i = 1 - p_i$ ($i = 2, 4, 6$). For additional details about canonical moments, see Dette and Studden (1997). Since ξ_s is a four-point design, $0 < p_2, p_4 < 1$. Thus (A.3) can be represented as

$$\det(\mathbf{M}(\xi_s)) = \frac{1}{(1 - \rho)^2(1 + \rho)^4} \times \phi(p_2, p_4, p_6),$$

where

$$\begin{aligned} \phi(p_2, p_4, p_6) &= p_2^3[2q_2p_4 + (1 + \rho)p_2] \\ &\quad \times \{q_2p_4[2q_4p_6 - (1 + \rho)(p_2p_4 - p_4 - 2p_2)] + (1 + \rho)p_2^2\}. \end{aligned}$$

In order to maximize $\det(\mathbf{M}(\xi_s))$, it is easy to see that we need $p_6 = 1$ and to maximize $\phi(p_2, p_4, 1)$ under the restriction $0 < p_2, p_4 < 1$. Therefore, p_2, p_4 must satisfy $(\partial/\partial p_2)\phi(p_2, p_4) = 0$ and $(\partial/\partial p_4)\phi(p_2, p_4) = 0$. Under the restriction that $0 < p_2, p_4 < 1$, for fixed p_2 , solving $(\partial/\partial p_4)\phi(p_2, p_4) = 0$ with respect to p_4 , it yields

$$p_4 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}, \tag{A.4}$$

where a, b, c are as defined in Theorem 3.2. Now substitute p_4 into $(\partial/\partial p_2)\phi(p_2, p_4) = 0$, then p_2 is a solution of Eq. (3.5) under the restriction that $0 < p_2 < 1$. Therefore, when $\rho < -\frac{2}{3}$, the D -optimal design $\xi_{0,\rho}^{(1,3)}$ is as stated in Theorem 3.2.

Proof of Theorem 5.1. Recall that the information matrix for model (2.1) with $r = m$ is

$$\mathbf{M}(\xi) = \frac{1}{1 - \rho^2} \begin{pmatrix} 2(1 - \rho)M_{11}(\xi) & (1 - \rho)M_{12}(\xi) & (1 - \rho)M_{12}(\xi) \\ \cdot & M_{22}(\xi) & -\rho M_{22}(\xi) \\ \cdot & \cdot & M_{22}(\xi) \end{pmatrix},$$

where $M_{ij}(\xi), 1 \leq i \leq j \leq 2$ is as defined in (3.1).

Simple calculation of the determinant of $\mathbf{M}(\xi)$ yields

$$\det(\mathbf{M}(\xi)) = (2(1 - \rho))^2 (1 - \rho^2)^{-(m+1)} (\det(M_{22}(\xi)))^2 \times \det(M_{11}(\xi) - M_{12}(\xi)M_{22}^{-1}(\xi)M_{12}^T(\xi)),$$

which indicates that the D -optimal design is independent of ρ . In the case of D_s -optimality, we need to maximize

$$\begin{aligned} & \det(\mathbf{M}(\xi)) / \det(2(1 - \rho)M_{11}(\xi)) \\ & = (1 - \rho^2)^{-(m+1)} \det(M_{22}(\xi)) \det(M_{22}(\xi) - M_{12}^T(\xi)M_{22}^{-1}(\xi)M_{12}(\xi)), \end{aligned}$$

which is also independent of ρ . \square

Proof of Theorem 5.2. Note that for $r = m = 2$, only the first four moments of design ξ appeared in the information matrix. Also by the symmetry argument of the D - and D_s -optimal designs and majorization argument of a matrix, the optimal designs should have $c_1 = c_3 = 0$, and $c_4 = c_2$. Then it is easy to see that the D - and D_s -optimal designs are supported only on $-1, 0, 1$ with corresponding weights as given in Theorem 5.2. \square

References

Abt, M., Gaffke, N., Liski, E.P., Sinha, B.K., 1998. Optimal designs in growth curve models: II Correlated model for quadratic growth: optimal designs for parameter estimation and growth prediction. J. Statist. Plann. Inference 67, 287–296.

- Abt, M., Liski, E.P., Mandal, N.K., Sinha, B.K., 1997. Optimal designs in growth curve models: I Correlated model for linear growth: Optimal designs for slope parameter estimation and growth prediction. *J. Statist. Plann. Inference* 64, 141–150.
- Atkins, J.E., Cheng, C.S., 1997. Optimal regression designs in the presence of random block effects. Technical Report. University of California, Berkeley.
- Bischoff, W., 1995. Determinant formulas with applications to designing when the observations are correlated. *Ann. Inst. Statist. Math.* 47, 385–399.
- Cheng, C.S., 1995. Optimal regression designs under random block-effects models. *Statist. Sinica* 5, 485–497.
- Dette, H., 1997a. Designing experiments with respect to ‘standardized’ optimality criteria. *J. Roy. Statist. Soc. B* 59, 97–110.
- Dette, H., 1997b. *E*-optimal designs for regression model with quantitative factors – a reasonable choice?. *Cand. J. Statist.* 25, 531–543.
- Dette, H., Studden, W.J., 1997. *The Theory of Canonical Moments with Applications in Statistics, Probability and Analysis*. Wiley, New York.
- Fedorov, V.V., 1972. *Theory of Optimal Experiments*. In: Studden, W.J., Kilmko, E.M. (Trans. and Ed.) Academic Press, New York.
- Graybill, F.A., 1983. *Matrices with Applications in Statistics*, 2nd Edition. Wadsworth International Group, Belmont, CA.
- Krafft, O., Schaefer, M., 1992. *D*-optimal designs for a multivariate regression model. *J. Multivariate Anal.* 42, 130–140.
- Pukelsheim, F., 1993. *Optimal Design of Experiments*. Wiley, New York.
- Studden, W.J., 1980. D_s -optimal designs for polynomial regression using continued fractions. *Ann. Statist.* 8, 1132–1141.