Statistica Sinica **21** (2011), 421-432

OPTIMALITY CRITERIA FOR MULTIRESPONSE LINEAR MODELS BASED ON PREDICTIVE ELLIPSOIDS

Xin Liu[1,2], Rong-Xian Yue[1,3] and Fred J. Hickernell[4]

[1]*Shanghai Normal University,* [2]*Donghua University,*
[3]*E-Institute of Shanghai Universities and* [4]*Illinois Institute of Technology*

Abstract: This paper proposes a new class of optimum design criteria for the linear regression model with r responses based on the volume of the predictive ellipsoid. This is referred to as I_L^r-optimality. The I_L^r-optimality criterion is invariant with respect to different parameterizations of the model, and reduces to I_L-optimality as proposed by Dette and O'Brien (1999) in single response situations. An equivalence theorem for I_L^r-optimality is provided and used to verify I_L^r-optimality of designs, and this is illustrated by several examples.

Key words and phrases: General equivalence theorem, multiresponse linear models, optimal design, predictive ellipsoid.

1. Introduction

In many experimental situations, especially in engineering, pharmaceutical, and biomedical, and environmental research, it is necessary to measure *more than one* response for each setting of control variables. For example, in the course of calibration of an apparatus in microwave engineering, several precision transmission line sections are connected to the apparatus. The connection of each section produces a complex number called a reflection coefficient; the reflection coefficients lie on a circle with unknown centre and radius, but due to various causes the readings are noisy. Berman (1983) analyzed such data using a simple bivariate response model. Wu (1997) considered Φ-optimal designs for Berman's model. Another example is a bioassay experiment that measures a response from different doses of the standard and test preparations. The expectation of the response at a dose level $d \in [a, b]$ under the standard preparation is $E(y_1|d) = \eta_1(d)$, while the expected response for the test preparation is $E(y_2|d) = \eta_2(d) = \eta_1(\tau d)$, where τ is a unknown constant representing the relative potency between the standard and test preparations. It is common practice to assume $\eta_1(d)$ is linearly related to $x = \log(d)$, and that the two responses are correlated. Huang et al. (2006) considered D-optimal designs for such models.

The previous work on multiresponse optimal designs has focused mainly on D-optimal designs. One of the earliest articles on multiresponse designs is the one by Draper and Hunter (1966) who developed a criterion for selecting additional experimental runs after a certain number of runs have already been chosen. Fedorov (1972, Chap. 5) established a theoretical foundation for multiresponse experiments and also developed a recursive algorithm for generating multiresponse approximate D-optimal designs. Chang (1994) studied the properties of D-optimal designs for multiresponse models. Khuri and Cornell (1996) devoted a chapter of their book to multiresponse experiments. Krafft and Schaefer (1992) considered a linear regression model with a one-dimensional control variable, and an m-dimensional response variable, and generated a D-optimal design for this special model. Chang et al. (2001) generated D-optimal designs for a simple m-dimensional response model with a single control variable.

Here we focus on the design criteria achieving reliable predictions from the fitted multiresponse linear models, since there are cases where prediction is also important when designing an experiment. We generalize the I_L-optimality of Dette and O'Brien (1999) for single response experiments to multiresponse situations. The I_L-optimality is analogous to Kiefer's Φ_k-criterion but is based on prediction variance, and contains G- and D-optimality as special cases.

The paper is organized as follows. Section 2 introduces the new optimality criteria for linear regression models with r responses, termed I_L^r-optimality. An equivalence theorem for I_L^r-optimality is derived in Section 3. Two illustrative examples are given in Section 4.

2. Development of I_L^r-Optimality

Throughout the paper we consider multiresponse linear models of the form

$$Y(x) = F(x)\theta + \varepsilon, \tag{2.1}$$

where $Y(x) = (y_1(x), \ldots, y_r(x))^T$ is the vector of r responses, $x = (x_1, \ldots, x_q)$ is a setting of q control variables, $F(x) = (f_1(x), \ldots, f_r(x))^T$ is an $r \times p$ matrix of regression functions, θ is a vector of p unknown parameters, and ε is a vector of random errors with mean 0 and known or unknown variance-covariance matrix Σ. We consider approximate designs of the form

$$\xi = \begin{Bmatrix} x_1 & \cdots & x_n \\ w_1 & \cdots & w_n \end{Bmatrix},$$

where the support points $x_1, \ldots, x_n$ are distinct elements of the design region $\mathcal{X} \subset \mathcal{R}^q$, and corresponding weights $w_1, \ldots, w_n$ are nonnegative real numbers which sum to unity. Denote the set of all approximate designs with positive semidefinite information matrix on $\mathcal{X}$ by Ξ. The information matrix of ξ is

$$M(\xi) = \int_{\mathcal{X}} F^T(x)\Sigma^{-1}F(x)d\xi(x),$$

and it is assumed that $\text{Range}\{F^T(x)\} \subset \text{Range}\{M(\xi)\}$ $(\forall x \in \mathcal{X})$, which implies that the r responses are estimable by the design ξ.

Motivated by Dette and O'Brien (1999), we take the matrix $F(x)$ of regression functions to be defined on a set $\mathcal{Z}$ that may be larger than the design region $\mathcal{X}$. It is assumed that $\mathcal{X}$ and $\mathcal{Z}$ are bounded, and μ denotes a probability measure on $\mathcal{Z}$.

For a point $z \in \mathcal{Z}$ the variance-covariance matrix of predicted responses at z under the design ξ is

$$V(z, \xi) = F(z)M^{-1}(\xi)F^T(z). \tag{2.2}$$

When there is no possibility of confusion we omit the dependencies of M, V and F on ξ and x.

Definition 1. For $L \in [1, \infty)$ a design ξ_L^* is called I_L^r-optimal in Ξ if it minimizes

$$\psi_L(\xi) = \left\{ \int_{\mathcal{Z}} |V(z, \xi)|^L \, d\mu(z) \right\}^{1/L} \tag{2.3}$$

over Ξ.

Remark 1. This definition can be extended to allow the case $L = \infty$ by taking $\psi_\infty(\xi) = \sup_{z \in \mathcal{Z}} |V(z, \xi)|$. It can be shown that $\psi_\infty(\xi) = \lim_{L \to \infty} \psi_L(\xi)$ if $\text{supp}(\mu) = \mathcal{Z}$ and each element of the matrix $F(z)$ of regression functions is continuous on $\mathcal{Z}$.

Obviously, the I_∞^r-optimality criterion, which minimizes the maximum volume of the predictive ellipsoid, is analogous to G-optimality in single response situations and can be viewed as a generation of G-optimality to multiresponse situations.

Remark 2. When there is only a single response, the determinant of the variance-covariance matrix, $|V(z, \xi)|$, degenerates to the variance function $d(z, \xi)$ and consequently the criterion function becomes

$$\psi_L(\xi) = \left\{ \int_{\mathcal{Z}} d^L(z, \xi) d\mu(z) \right\}^{1/L},$$

which is the I_L-optimality criterion introduced by Dette and O'Brien (1999) in single response situations. On the other hand, if $r = p$, and hence $F(x)$ is an $p \times p$ matrix, for example, $\eta_1(x, \theta) = \theta_1 + \theta_2 x$, $\eta_2(x, \theta) = \theta_1 + \theta_2 e^x$, $0 \le x \le 1$, then

$$|V(z, \xi)| = |F(z)|^2 \left| M^{-1}(\xi) \right|,$$

and I_L^r-optimality is equivalent to D-optimality.

Comparing with Kiefer's Φ_k class, a good property of I_L^r-optimality is that it is invariant with respect to model reparameterization. Thus the matrix of regression functions $F(x)$ can be replaced by $G(x) := F(x)A$ for any nonsingular $p\times p$ matrix A and θ replaced by $\gamma := A^{-1}\theta$. This was also noted for I_L-optimality by Dette and O'Brien (1999, Thm. 1).

3. An equivalence theorem for I_L^r-optimality

It is well known that the general equivalence theorem plays an important role in optimal approximate design theory. Here we establish an equivalence theorem for I_L^r-optimality to characterize I_L^r-optimal designs.

Lemma 1. *Let $\mathcal{P}_n$ denote the set of all positive definite matrices of order $n \times n$ and suppose A is a fixed $m \times n (n \geq m)$ matrix. Then $f(B) = |AB^{-1}A^T|$ is convex on $\mathcal{P}_n$.*

This lemma is a special case of results in Gaffke and Heiligers (1996, p.1153). From this the lemma and (2.3), we have the following.

Lemma 2. *For the criterion function ψ_L defined by* (2.3) *we have:*

(i) *ψ_L is convex on Ξ;*

(ii) *the directional derivative of ψ_L at ξ in the direction of $\bar{\xi}$, denoted $\Delta_{\psi_L}(\xi,\bar{\xi})$, is*

$$\Delta_{\psi_L}(\xi,\bar{\xi}) = r\psi_L(\xi) - \psi_L^{1-L}(\xi)\mathrm{tr}\left\{M^{-1}(\xi)M(\bar{\xi})M^{-1}(\xi)\int_{\mathcal{Z}}|V|^L F^T V^{-1} F d\mu\right\};$$

(iii) *for any fixed ξ with nonsingular information matrix, the directional derivative*

$$\Delta_{\psi_L}(\xi,\bar{\xi}) = \int_{\mathcal{X}} \Delta_{\psi_L}(\xi,\delta_x)d\bar{\xi}(x), \quad (\xi,\bar{\xi}) \in \Xi \times \Xi,$$

is linear in $\bar{\xi}$, where $\delta_x \in \Xi$ denotes the Dirac measure at x.

Proof. (i) The convexity of ψ_L follows immediately from Lemma 1 and Minkowski's inequality.

(ii) Let $\xi, \bar{\xi} \in \Xi,\ \alpha \in (0,1)$ and $\xi_\alpha = (1-\alpha)\xi + \alpha\bar{\xi}$. We have

$$\begin{aligned}\frac{d}{d\alpha}|V(z,\xi_\alpha)| &= |V(z,\xi_\alpha)|\,\mathrm{tr}\left\{V^{-1}(z,\xi_\alpha)\frac{d}{d\alpha}\,V(z,\xi_\alpha)\right\}\\ &= |V(z,\xi_\alpha)|\,\mathrm{tr}\left\{V^{-1}(z,\xi_\alpha)F(z)M^{-1}(\xi_\alpha)\left(M(\xi)-M(\bar{\xi})\right)M^{-1}(\xi_\alpha)F^T(z)\right\},\end{aligned}$$

so that for all $L \in [1, \infty)$,

$$\begin{aligned}\Delta_{\psi_L}(\xi, \bar{\xi}) &= \lim_{\alpha \to 0^+} \frac{d\psi_L(\xi_\alpha)}{d\alpha} \\ &= \psi_L^{1-L}(\xi) \int_{\mathcal{Z}} \left(|V|^L \operatorname{tr}\left\{ V^{-1}\left(V - F(z)M^{-1}(\xi)M(\bar{\xi})M^{-1}(\xi)F^T(z)\right)\right\}\right) d\mu(z) \\ &= r\psi_L(\xi) - \psi_L^{1-L}(\xi)\operatorname{tr}\left\{ M^{-1}(\xi)M(\bar{\xi})M^{-1}(\xi)\int_{\mathcal{Z}} |V|^L F^T(z)V^{-1}F(z)d\mu(z)\right\}.\end{aligned}$$

(iii) The linearity of $\Delta_{\psi_L}(\xi, \bar{\xi})$ in $\bar{\xi}$ can be obtained by noting that $M(\bar{\xi}) = \int_{\mathcal{X}} M(\delta_x)d\bar{\xi}(x)$, and the proof is complete.

According to Whittle (1973), ξ_L^* is I_L^r-optimal in $\Xi \iff \inf_{x\in\mathcal{X}} \Delta_{\psi_L}(\xi_L^*, \delta_x) = 0$, which implies the following.

Theorem 1. *For all $L \in [1, \infty)$, a design $\xi_L^* \in \Xi$ is I_L^r-optimal in Ξ if and only if*

$$\begin{aligned}&\sup_{x\in\mathcal{X}} tr\Big\{ M^{-1}(\xi_L^*)F^T(x)\Sigma^{-1}F(x)M^{-1}(\xi_L^*)\int_{\mathcal{Z}}|V(z,\xi_L^*)|^L F^T(z)V^{-1}(z,\xi_L^*)F(z)d\mu(z)\Big\} \\ &= r\int_{\mathcal{Z}} |V(z,\xi_L^*)|^L d\mu(z). \qquad (3.1)\end{aligned}$$

Moreover, the supremum is achieved at the support points of ξ_L^*.

In order to compare the performance of different designs, e.g., I_1^r- and I_∞^r-optimal designs, we define the efficiency of a design ξ as

$$\text{Eff}_L(\xi) = \frac{\psi_L(\xi_L^*)}{\psi_L(\xi)}, \qquad (3.2)$$

where ξ_L^* denotes the I_L^r-optimal design. It is clear that $0 \leqslant \text{Eff}_L(\xi) \leqslant 1$ for all $\xi \in \Xi$. The following corollary provides a lower bound for $\text{Eff}_L(\xi)$ that follows immediately from Theorem 1 and Pilz (1983, p.137, Lemma 11.5).

Corollary 1. *For $L \in [0, \infty)$, if*

$$\phi_L(x,\xi) = \frac{tr\Big\{M^{-1}(\xi)F^T(x)\Sigma^{-1}F(x)M^{-1}(\xi)\int_{\mathcal{Z}}|V(z,\xi)|^L F^T(z)V^{-1}(z,\xi)F(z)d\mu(z)\Big\}}{\int_{\mathcal{Z}}|V(z,\xi)|^L d\mu(z)}, \qquad (3.3)$$

then $\mathit{Eff}_L(\xi) \geq 1 + r - \sup_{x\in\mathcal{X}} \phi_L(x,\xi)$.

In terms of the function $\phi_L(x,\xi)$ at (3.3), we can restate Theorem 1.

Theorem 1′. *For $L \in [0, \infty)$, a design $\xi_L^* \in \Xi$ is I_L^r-optimal in Ξ if and only if $\sup_{x\in\mathcal{X}} \phi_L(x, \xi_L^*) = r$. Moreover, the supremum is achieved at the support points of ξ_L^*.*

According to Dette and O'Brien (1999), the equivalence theorem can be extended to the case $L = \infty$. For any $\xi \in \Xi$ we define the answering set (Danskin (1967, p.21))

$$\mathcal{A}(\xi) = \left\{ z \in \mathcal{Z} \,\middle|\, |V(z,\xi)| = \sup_{z'\in\mathcal{Z}} \left|V(z',\xi)\right| \right\}.$$

Let μ^* be a probability measure on $\mathcal{A}(\xi)$ and define

$$\phi_\infty(x,\xi) = \mathrm{tr}\left\{ M^{-1}(\xi)F^T(x)\Sigma^{-1}F(x)M^{-1}(\xi)\int_{\mathcal{A}(\xi)} F^T(z)V^{-1}(z,\xi)F(z)d\mu^*(z) \right\}. \tag{3.4}$$

Theorem 2. *A design $\xi_\infty^* \in \Xi$ is I_∞^r-optimal in Ξ if and only if there exists a probability measure μ^* on $\mathcal{A}(\xi_\infty^*)$ such that*

$$\sup_{x\in\mathcal{X}} \phi_\infty(x, \xi_\infty^*) = r.$$

Moreover, the supremum is achieved at the support points of ξ_∞^*.

4. Illustrative Examples

In this section, we present two examples of I_1^r- and I_∞^r-optimal designs for bivariate response models. In Example 1, for a linear and quadratic regression model, we state designs and prove their optimality by means of the equivalence theorems. In Example 2 we consider Berman's (1983) model, and construct I_1^r- and I_∞^r-optimal designs immediately.

Example 1. Linear and Quadratic regression. For $\mathcal{X} = \mathcal{Z} = [0,1]$, we consider the responses

$$\begin{cases} \eta_1(x,\theta_1) = \theta_{10} + \theta_{11}x, \\ \eta_2(x,\theta_2) = \theta_{20} + \theta_{21}x + \theta_{22}x^2. \end{cases} \tag{4.1}$$

Let Σ be the variance-covariance matrix of the response vector. We take $d\mu(z) = dz$. For this model, the matrix of regression functions is

$$F(x) = \begin{pmatrix} 1 & x & 0 & 0 & 0 \\ 0 & 0 & 1 & x & x^2 \end{pmatrix},$$

and the vector of parameters is $\theta = (\theta_{10}, \theta_{11}, \theta_{20}, \theta_{21}, \theta_{22})^T$. The D-optimal design for this model was found by Krafft and Schaefer (1992) as follows:

$$\xi_D^* = \left\{ \begin{array}{ccc} 0 & \frac{1}{2} & 1 \\ w_D & 1-2w_D & w_D \end{array} \right\}, \quad \text{where } w_D = \frac{3}{8}. \tag{4.2}$$

Note that the support points of ξ_D^* are the endpoints and the centre point of the design space $\mathcal{X} = [0, 1]$, and the weights on the two endpoints are equal.

Motivated by the D-optimal design and admissibility for the second-order polynomial regression (Pukelsheim (1993, p.253)), we guessed the I_L^r-optimal design to be

$$\xi^* = \left\{ \begin{array}{ccc} 0 & \frac{1}{2} & 1 \\ w & 1-2w & w \end{array} \right\},$$

where $w \in [0, 1/2]$ is an unknown parameter to be determined. We claim that the I_1^r- and I_∞^r-optimal designs are given by

$$\xi_1^* = \left\{ \begin{array}{ccc} 0 & \frac{1}{2} & 1 \\ w_1 & 1-2w_1 & w_1 \end{array} \right\}, \quad \text{where } w_1 = \frac{2\sqrt{22}-5}{14}, \tag{4.3}$$

$$\xi_\infty^* = \left\{ \begin{array}{ccc} 0 & \frac{1}{2} & 1 \\ w_\infty & 1-2w_\infty & w_\infty \end{array} \right\}, \quad \text{where } w_\infty = \frac{\sqrt{6}}{6}. \tag{4.4}$$

We now verify optimality by Theorems $1'$ and 2, respectively. For I_1^r-optimality of ξ_1^*, straightforward calculation gives the following expressions for the determinant of the matrix $V(x, \xi_1^*)$ defined by (2.2) and the function $\phi_L(x, \xi_1^*)$ defined by (3.3) for any $x \in [0, 1]$:

$$|V(x, \xi_1^*)| = \frac{|\Sigma|(1+2w_1-4x+4x^2)(1-2w_1+12w_1x-12w_1x^2-6x+14x^2-16x^3+8x^4)}{2w_1^2(1-2w_1)},$$

where $w_1 = (2\sqrt{22}-5)/14$, and

$$\phi_1(x, \xi_1^*) = 2 - \frac{2(170-49\sqrt{22})x(1-x)(2x-1)^2}{27}.$$

It is clear that $\phi_1(x, \xi_1^*)$ is nonnegative for any $x \in [0, 1]$, and attains its maximum $r = 2$ at $x \in \{0, 1/2, 1\}$, the support points of ξ_1^*. It follows from Theorem $1'$ that the design ξ_1^* is I_1^r-optimal over the class Ξ.

For I_∞^r-optimality of ξ_∞^*, straightforward calculation gives the determinant of $V(x, \xi_\infty^*)$ as

$$|V(x, \xi_\infty^*)| = \frac{|\Sigma|(1+2w_\infty-4x+4x^2)(1-2w_\infty+12w_\infty x-12w_\infty x^2-6x+14x^2-16x^3+8x^4)}{2w_\infty^2(1-2w_\infty)},$$

where $w_\infty = \sqrt{6}/6$. The answering set corresponding to ξ_∞^* is

$$\mathcal{A}(\xi_\infty^*) = \left\{ z \in \mathcal{Z} \,\middle|\, |V(z, \xi_\infty^*)| = \sup_{z' \in \mathcal{Z}} \left| V(z', \xi_\infty^*) \right| \right\} = \left\{ 0,\ \frac{1}{2},\ 1 \right\}.$$

Table 1. Efficiencies of the multiresponse designs $\xi_1^*, \xi_\infty^*, \xi_D^*$ and the single response designs $\zeta_1^*, \zeta_\infty^*$ for the multiresponse model in (4.1)

ξ	ξ_1^*	ξ_∞^*	ξ_D^*	ζ_1^*	ζ_∞^*
$\mathrm{Eff}_1(\xi)$	1.0000	0.7591	0.9010	0.9131	0.9898
$\mathrm{Eff}_\infty(\xi)$	0.6564	1.0000	0.8758	0.4541	0.7266

We take the probability measure μ^* corresponding to I_∞^r-optimality as

$$\mu^* = \left\{ \begin{array}{ccc} 0 & \frac{1}{2} & 1 \\ s & 1-2s & s \end{array} \right\}, \quad \text{where} \quad s = \frac{2+\sqrt{6}}{12}.$$

Straightforward calculation gives $\phi_\infty(x, \xi_\infty^*) = 2-(5+\sqrt{6})x(1-x)(2x-1)^2$. It is clear that $\phi_\infty(x, \xi_\infty^*)$ is nonnegative for any $x \in [0,1]$, and attains its maximum $r = 2$ at $x \in \{0, 1/2, 1\}$, the support points of ξ_∞^*. It follows from Theorem 2 that the design ξ_∞^* is I_∞^r-optimal over the class Ξ.

Figure 1 shows the graphs of $|V(z, \xi^*)|$ corresponding to the D-, I_1^r-, and I_∞^r-optimal designs given in (4.2), (4.3) and (4.4), respectively, indicating how the optimal designs weight the regions of the prediction space differently; it is assumed that $|\Sigma| = 1$ without loss of generality. Observe that $|V(z, \xi_1^*)|$ lies below both $|V(z, \xi_\infty^*)|$ and $|V(z, \xi_D^*)|$ for about three-fourths of the prediction space. The efficiencies Eff_1 and Eff_∞ of ξ_1^*, ξ_∞^* and ξ_D^* for model (4.1) are given in Table 1. One sees that the performance of the D-optimal design ξ_D^* compares well with both ξ_1^* and ξ_∞^*.

For comparison, we also consider the performance of the I_1- and I_∞-optimal designs for each single response against the I_1^r- and I_∞^r-optimal designs for the multiresponse model in (4.1). Note that the I_1- and I_∞-optimal designs for the individual response $\eta_1(x, \theta_1) = \theta_{10} + \theta_{11}x$ on $\mathcal{X} = \mathcal{Z} = [0,1]$ has two support points 0 and 1, and consequently is singular for the multiresponse model. The I_1- and I_∞-optimal designs for the individual response $\eta_2(x, \theta_2) = \theta_{20} + \theta_{21}x + \theta_{22}x^2$ on $\mathcal{X} = \mathcal{Z} = [0,1]$, denoted by ζ_1^* and ζ_2^*, respectively, are as follows (Dette and O'Brien (1999)):

$$\zeta_1^* = \left\{ \begin{array}{ccc} 0 & \frac{1}{2} & 1 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{array} \right\}, \quad \zeta_\infty^* = \left\{ \begin{array}{ccc} 0 & \frac{1}{2} & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \right\}.$$

The efficiencies Eff_1 and Eff_∞ of ζ_1^* and ζ_∞^* are given in Table 1. None that ζ_∞^* is nearly I_1^r-optimal for the multiresponse model, and the performance of ζ_1^* also compares very well with ξ_1^*. On the other hand, the performance of ζ_∞^* is moderate, and ζ_1^* is poor compared with the design ξ_∞^* for the multiresponse model.

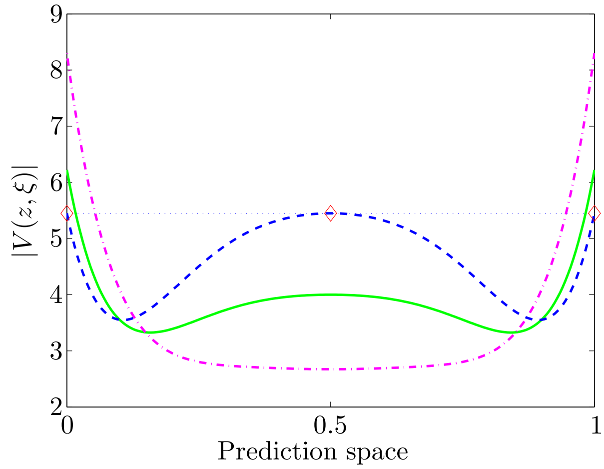

Figure 1. The graphs of $|V(z,\xi)|$ for Example 1 with $|\Sigma| = 1$: $\xi = \xi_D^*$ (solid line), $\xi = \xi_1^*$ (dash-dot line) and $\xi = \xi_\infty^*$ (dashed line).

Example 2.(Berman's (1983) model.) Wu (1997) represents the Berman's model on a circular arc $\mathcal{X} = [-\alpha/2, \alpha/2]$ for an arc of length $\alpha \in (0, 2\pi]$ by the bivariate four-parameter linear model

$$\begin{cases} \eta_1(t,\theta) = \theta_1 + \theta_3 \cos t - \theta_4 \sin t, \\ \eta_2(t,\theta) = \theta_2 + \theta_3 \sin t + \theta_4 \cos t, \end{cases} \qquad t \in \mathcal{X} = [-\frac{\alpha}{2}, \frac{\alpha}{2}], \tag{4.5}$$

or, by $Y(t) = F(t)\theta + \varepsilon$, where the matrix of regression functions is

$$F(t) = (I_2, A(t)), \quad \text{where } A(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

and $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)^T$. The variance-covariance matrix of $\varepsilon = (\varepsilon_1, \varepsilon_2)^T$ is assumed to be $\Sigma = \sigma^2 I_2$. In what follows we can easily construct I_1^r- and I_∞^r-optimal designs for this model. It is assumed that $\mathcal{Z} = \mathcal{X}$ and $d\mu(z) = \alpha^{-1}dz$.

For a design ξ, the information matrix is

$$M(\xi) = \begin{pmatrix} I_2 & A(\xi) \\ A^T(\xi) & I_2 \end{pmatrix},$$

where

$$A(\xi) = \int_{\mathcal{X}} A(t)\, d\xi = \begin{pmatrix} c(\xi) & -s(\xi) \\ s(\xi) & c(\xi) \end{pmatrix}, \; c(\xi) = \int_{\mathcal{X}} \cos t\, d\xi, \quad s(\xi) = \int_{\mathcal{X}} \sin t\, d\xi.$$

Note that if $\tilde{\xi}$ denotes the reflection of a design ξ across the midpoint of the arc, then $\psi_L(\xi) = \psi_L(\tilde{\xi})$. Consequently, if ξ is I_L^r-optimal, then $\tilde{\xi}$ is I_L^r-optimal,

Table 2. Values of t^* in (4.6) for various α in $(0, 2\pi]$.

α	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3\pi}{4}$	π	$\frac{5\pi}{4}$	$\frac{3\pi}{2}$	$\frac{7\pi}{4}$	2π
t^*	$\frac{\pi}{8}$	$\frac{\pi}{4}$	1.0931	1.2333	1.3466	1.4389	1.5131	$\frac{\pi}{2}$

and the symmetrized design $\xi^* = (\xi + \tilde{\xi})/2$ is also I_L^r-optimal. Therefore, it is sufficient to search for I_L^r-optimal designs among the symmetric designs on $\mathcal{X} = [-\alpha/2, \alpha/2]$.

For a symmetric design ξ we have $s(\xi) = 0$ and

$$|V(t,\xi)| = 4\left(\frac{1 - c(\xi)\cos t}{1 - c^2(\xi)}\right)^2,$$

which depends on ξ only through the cos-term $c(\xi)$.

For $L = 1$,

$$\psi_1(\xi) = g(c(\xi)) := \frac{2}{\alpha(1 - c^2(\xi))^2}\left[2\alpha - 8c(\xi)\sin\frac{\alpha}{2} + c^2(\xi)(\alpha + \sin\alpha)\right].$$

Let c^* be a minimizer of $g(c)$. Then a design ξ^* that satisfies the equation $c(\xi^*) = c^*$ is I_1^r-optimal. It is not difficult to verify that the following is a I_1^r-optimal design:

$$\xi_{1,\alpha}^* = \begin{Bmatrix} -t^* & t^* \\ \frac{1}{2} & \frac{1}{2} \end{Bmatrix}, \qquad \text{where} \quad t^* = \arccos(c^*). \tag{4.6}$$

Note that c^* depends on α. Table 2 shows the values of t^* in (4.6) for various $\alpha \in (0, 2\pi]$.

For $L = \infty$, a straightforward argument gives that

$$\psi_\infty(\xi) = h(c(\xi)) := \begin{cases} 4\left(\dfrac{1 - c(\xi)\cos(\alpha/2)}{1 - c^2(\xi)}\right)^2, & \text{if } 0 \le c(\xi) < 1, \\ \dfrac{4}{(1 + c(\xi))^2}, & \text{if } \cos\frac{\alpha}{2} \le c(\xi) < 0. \end{cases}$$

For $0 < \alpha < \pi$, the function $h(c)$ is minimized at $c^* = \cos(\alpha/2)$, and then the following design is I_∞^r-optimal:

$$\xi_{\infty,\alpha}^* = \begin{Bmatrix} -\frac{\alpha}{2} & \frac{\alpha}{2} \\ \frac{1}{2} & \frac{1}{2} \end{Bmatrix}. \tag{4.7}$$

For $\pi \le \alpha \le 2\pi$, $h(c)$ is minimized at $c^* = 0$, and then the design ξ^* which satisfies $c(\xi^*) = 0$ is I_∞^r-optimal. A I_∞^r-optimal design is

$$\xi_{\infty,\alpha}^* = \begin{Bmatrix} -\frac{\pi}{2} & \frac{\pi}{2} \\ \frac{1}{2} & \frac{1}{2} \end{Bmatrix}. \tag{4.8}$$

Note that the design $\xi^*_{\infty,\alpha}$ ($\pi \leq \alpha \leq 2\pi$) in (4.8) is also an orthogonal design for model (4.5), i.e., the information matrix of $\xi^*_{\infty,\alpha}$ is diagonal.

Acknowledgements

This work is partially supported by NSFC grant (11071168), E-Institutes of Shanghai Municipal Education Commission (E03004), Shanghai Leading Academic Discipline Project (S30405), the Innovation Program of Shanghai Municipal Education Commission (11zz116) and the Special Funds for Doctoral Authorities of Education Ministry. The authors would like to thank the reviewers and editors for their constructive comments on an earlier version of this paper.

References

Berman, M. (1983). Estimating the parameters of a circle when angular differences are known. *J. Roy. Statist. Soc. Ser. C* **32**, 1-6.

Chang, S. I. (1994). Some properties of multi-response *D*-optimal designs. *J. Math. Anal. Appl.* **184**, 256-262.

Chang, F.-C., Huang, M.-N. L., Lin, D. K. J. and Yang H.-C. (2001). Optimal designs for dual response polynomial regression models. *J. Statist. Plann. Inference* **93**, 309-322.

Danskin, J. M. (1967). *The Theory of Max-min and its Application to Weapons Allocation Problems.* Springer-Verlag, New York.

Dette, H. and O'Brien, T. E. (1999). Optimality criteria for regression models based on predicted variance. *Biometrika* **86**, 93-106.

Draper, N. R. and Hunter, W. G. (1966). Design of experiments for parameter estimation in multiresponse situations. *Biometrika* **53**, 525-533.

Fedorov, V. V. (1972). *Theory of Optimal Experiments.* Academic Press, New York.

Gaffke, H. and Heiligers, B. (1996). Approximate designs for polynomial regression: Invariance, admissibility, and optimality. In Handbook of Statistics: *Design and Analysis of Experiments* (Edited by S. Ghosh and C. R. Rao), 1149-1199. Elsevier, Amsterdam.

Huang, M.-N. L., Chen, R. B., Lin, C. S. and Wong, W. K. (2006). Optimal designs for parallel models with correlated responses. *Statist. Sinica* **16**, 121-133.

Khuri, A. I. and Cornell, J. A. (1996). *Response Surfaces: Designs and Analysis.* 2nd. Edn. Marcel Dekker, New York.

Krafft, O. and Schaefer, M. (1992). D-optimal designs for a multivariate regression model. *J. Multivariate Anal.* **42**, 130-140.

Pilz, J. (1983). *Bayesian Estimation and Experimental Design in Linear Regression Models.* Teubner, Leipzig.

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

Whittle, P. (1973). Some general points in the theory of optimal experimental designs. *J. Roy. Statist. Soc. Ser. B* **35**, 123-130.

Wu, H. (1997). Optimal exact designs on a circle or a circular arc. *Ann. Statist.* **25**, 2027-2043.

College of Science, Donghua University, Shanghai 201600, P. R. China.

E-mail: cxqs2006@yahoo.com.cn

Department of Mathematics, Shanghai Normal University, E-Institute of Shanghai Universities, and Scientific Computing Key Laboratory of Shanghai Universities, Shanghai 200234, P. R. China.

E-mail: yue2@shnu.edu.cn

Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL 60616-3793, U.S.A.

E-mail: hickernell@iit.edu

(Received November 2008; accepted August 2009)