Statistica Sinica **29** (2019), 2187-2203
doi:https://doi.org/10.5705/ss.202017.0332

# DESIGN ADMISSIBILITY, INVARIANCE, AND OPTIMALITY IN MULTIRESPONSE LINEAR MODELS

Xin Liu and Rong-Xian Yue

*Donghua University and Shanghai Normal University*

*Abstract:* This study examines optimal design problems in multiresponse linear models. We investigate the optimality, admissibility, and invariance of approximate designs. Necessary and sufficient conditions are given for a design to be admissible and invariant. Elfving's theorem for $D$-optimality is established for the multiresponse linear models.

*Key words and phrases:* Admissible design, Elfving's theorem, invariant design, multiresponse model, optimal design.

## 1. Introduction

Optimal experimental designs have been applied successfully in many areas, including engineering, biomedical, environmental, and epidemiological research, since the seminal work of Smith (1918). Identifying an optimal design often results in an intricate optimization problem that is difficult to handle. In the field of optimal designs, current available tools are based mainly on the general equivalence theorem of Kiefer and Wolfowitz (1959) and the geometric approach of Elfving (1952). The general equivalence theorem provides the necessary and sufficient conditions for a design to be optimal under a specific criterion. This provides a way to check the optimality of a candidate design and to construct optimal designs iteratively. The geometric approach presents a geometric characterization of optimal designs and allows us to search for designs with support included only in the "extreme points" of the Elfving set. See Silvey and Titterington (1973), Dette (1993), Dette and Studden (1993), Dette and Holland-Letz (2009), and Holland-Letz, Dette and Pepelyshev (2011), for example. Because of the complicated structure of the corresponding optimization problems, general results are extremely difficult to obtain. This means that results can only be obtained on a case-by-case basis.

A useful strategy is to simplify the design problem by identifying a complete subclass, $\Xi_{com}$, composed of relatively simple designs. In addition, the subclass is sufficiently small that for any design $\xi$ not belonging to this class, there is a design in the class that has an information matrix dominating that of $\xi$ in the Loewner ordering. We may then restrict our attention to this subclass $\Xi_{com}$. Along this line, a series of remarkable papers by Yang and Stufken (2009, 2012); Yang (2010); Dette and Melas (2011), and Dette and Schorning (2013) derived several complete classes of designs for single-response models with respect to the Loewner ordering of the information matrices, based on considerations of admissibility and invariance.

In many experimental situations, especially in engineering, pharmaceutical, biomedical, and environmental research, more than one response is measured for each unit. Thus, multiresponse models play an important role in many areas of science. For example, the model of Berman (1983) is used to analyze data obtained when calibrating apparatus in microwave engineering. Another example is a bioassay experiment that measures a response from different doses of standard and test preparations, which can be fitted by the parallel linear model introduced in Huang et al. (2006). For more examples of multiresponse statistical models, refer to Atkinson and Bogacka (2002) and Uciéski and Bogacka (2005). Although work on the theory of an optimal design for single-response models dates back as far as 1918, multiresponse models did not appear in the optimal design literature until 1966. Draper and Hunter (1966) developed a criterion for selecting additional experiment runs after a certain number of runs have already been chosen. However, the literature in this area is relatively sparse owing to the increased theoretical and computational challenges associated with multiresponse models. For further reference, refer to Khuri (1990, 1996) and Liu, Yue and Hickernell (2011), and the literature cited therein.

As mentioned above, considerations of admissibility and invariance are key to reducing complicated design problems, which have been applied successfully to finding optimal designs by many authors, including Kiefer and Wolfowitz (1959). However, while these techniques have been discussed in detail in the context of single-response models (see Pukelsheim (1993)), they are underdeveloped for multiresponse models.

In the present study, we consider admissibility and invariance in the design problem for multiresponse linear models. Our strategy is to reformulate the multiresponse linear model in a simple form so that the original design problem can be transformed into an equivalent problem for a corresponding single-response model.

The rest of the paper is organized as follows. In Section 2, we first specify the multiresponse linear model, and then provide reformulations of the multiresponse linear model and its information matrix. Three examples of multiresponse models are also given in this section. Sections 3 and 4 consider design admissibility and invariance, respectively. In Section 5, we establish Elfving's theorem for $D$-optimality in the multiresponse linear models. Section 6 concludes the paper.

## 2. Model Specification and Reformulation

### 2.1. Model specification

We consider the following multiresponse linear model:

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

where $Y(x) = (y_1(x), \ldots, y_r(x))^T$ is an $r$-dimensional response, $x = (x_1, \ldots, x_q)$ 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, $\theta$ is a vector of $p$ unknown parameters, and $\varepsilon$ is an $r$-dimensional vector of random errors, with mean zero and nonsingular covariance matrix $\Sigma = (\sigma_{ij})_{r\times r}$. The following are three examples of multiresponse linear models.

**Example 1.** Linear and quadratic model (see Krafft and Schaefer (1992)).

The linear and quadratic model on $\mathcal{X} = [-1, 1]$ is as follows:

$$\begin{cases} y_1(x) = \theta_{10} + \theta_{11}x + \varepsilon_1, \\ y_2(x) = \theta_{20} + \theta_{21}x + \theta_{22}x^2 + \varepsilon_2, \end{cases} \tag{2.2}$$

or simply by (2.1), where

$$F(x) = \begin{pmatrix} 1 & x & 0 & 0 & 0 \\ 0 & 0 & 1 & x & x^2 \end{pmatrix}, \quad \theta = (\theta_{10}, \theta_{11}, \theta_{20}, \theta_{21}, \theta_{22})^T.$$

Models of this type are frequently used to describe chemical reactions, where $x$ may represent time or temperature and $y_i$ describes concentrations of the various substances involved. Consider, by way of illustration, an experiment on the decomposition of aspartame, a synthetic sweetener (see (Soo and Bates, 1996, Sec. 5)). The principal product of the decomposition of aspartame (APM) is diketopiperazine (DKP). Model (2.2), with $x \in [a, b] = [0, 5]$, can be used to describe the concentrations of APM and DKP during the first five seconds of the experiment.

**Example 2.** The Berman model in Berman (1983).

The Berman model on a circular arc $\mathcal{X} = [-\alpha/2, \alpha/2]$, for an arc of length $\alpha \in [0, 2\pi]$, is represented by

$$\begin{cases} y_1(t) = \theta_1 + \theta_3 \cos t - \theta_4 \sin t + \varepsilon_1, \\ y_2(t) = \theta_2 + \theta_3 \sin t + \theta_4 \cos t + \varepsilon_2, \end{cases} \tag{2.3}$$

or simply by (2.1), where

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

and $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)^T$. The covariance matrix of $\varepsilon = (\varepsilon_1, \varepsilon_2)^T$ is assumed to be $\Sigma = \sigma^2 I_2$.

This model was proposed by Berman (1983) with two particular applications: (i) the calibration of an impedance measuring apparatus in microwave engineering, and (ii) the analysis of megalithic sites in Britain, in which archaeologists need to fit circles to stone rings. Model (2.3) assumes that the angular differences between sample points are known in advance, either from the special structure of the problem or through experimental design. It is a frequently used model for fitting circular data.

**Example 3.** Parallel linear model with two responses (see Huang et al. (2006)).

The parallel linear model on $\mathcal{X} = [-1, 1]^2$ is described by

$$\begin{cases} y_1(x) = \theta_{01} + \theta_1 x_1 + \varepsilon_1, \\ y_2(x) = \theta_{02} + \theta_1 x_2 + \varepsilon_2. \end{cases} \tag{2.4}$$

The covariance matrix of $\varepsilon = (\varepsilon_1, \varepsilon_2)^T$ is assumed to be $\Sigma = (1-\rho)I_2 + \rho J_2$, where $I_2$ is the identity matrix of order 2 and $J_2$ is an all-ones matrix of order 2.

In this model, $r = 2$, $p = 3$, and

$$F(x) = \begin{pmatrix} 1 & 0 & x_1 \\ 0 & 1 & x_2 \end{pmatrix}, \quad \theta = (\theta_{01}, \theta_{02}, \theta_1)^T.$$

This model can be used in analyses of bioassay experiments to measure responses 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 $E(y_1|d) = \eta_1(d)$. The expected response for the test preparation is $E(y_2|d) = \eta_2(d) = \eta_1(\tau d)$, where $\tau$ is an 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.

This study investigates approximate designs that are probability measures on the design region $\mathcal{X}$ with finite support, which we denote by

$$\xi = \begin{Bmatrix} x_1 & x_2 & \dots & x_n \\ w_1 & w_2 & \dots & w_n \end{Bmatrix}, \quad 0 < w_i \le 1, \quad \sum_{i=1}^{n} w_i = 1.$$

Here, $x_i$ denotes a support point at which a measurement is taken, and $w_i$ is the weight assigned to each level in the design. The information matrix of a design $\xi$ for the model (2.1) is given by

$$M(\xi) = \int_{\mathcal{X}} F^T(x)\Sigma^{-1}F(x)d\xi(x). \tag{2.5}$$

We use the notation $\Xi$ for the set of all approximate designs, and $\mathcal{M}(\Xi)$ for the set of all information matrices on $\Xi$. It is assumed that $\text{Ran}(F^T(x)) \subset \text{Ran}(M(\xi))$, which implies that the $r$ responses are estimable by the design $\xi$, where $\text{Ran}(A)$ denotes the range of matrix $A$.

## 2.2. Model reformulation

In this subsection, we reformulate the multiresponse model (2.1) and its information matrix (2.5), which are used in the following sections.

Let $f(x) = (l_1(x), \dots, l_k(x))^T$ be a vector consisting of all different elements in $F(x)$, where $k$ is the total number of different elements in $F(x)$. Then the $i$-th regression vector is denoted as $f_i(x) = V_i^T U_i f(x)$, where $U_i$ and $V_i$, for $i = 1, \dots, r$, are full row-rank matrices satisfying $f_i^T(x)\theta = f^T(x)U_i^T V_i \theta$, for $i = 1, \dots, r$. Then, the $r \times p$ matrix $F(x)$ in (2.1) can be rewritten as

$$F(x) = (f_1(x), \dots, f_r(x))^T = (V_1^T U_1 f(x), \dots, V_r^T U_r f(x))^T \tag{2.6}$$

$$= \begin{pmatrix} f^T & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & f^T \end{pmatrix} \begin{pmatrix} U_1^T V_1 \\ \vdots \\ U_r^T V_r \end{pmatrix} = [I_r \otimes f^T(x)] L_{UV}, \tag{2.7}$$

where $L_{UV} = (V_1^T U_1, \dots, V_r^T U_r)^T$. Consequently, model (2.1) can be rewritten in the following form:

$$Y(x) = [I_r \otimes f^T(x)] L_{UV}\theta + \varepsilon. \tag{2.8}$$

Correspondingly, the information matrix (2.5) of design $\xi$ is expressed as follows:

$$\begin{aligned} M(\xi) &= \int_{\mathcal{X}} ([I_r \otimes f^T(x)]L_{UV})^T \Sigma^{-1} [I_r \otimes f^T(x)] L_{UV} d\xi(x) \qquad (2.9) \\ &= \int_{\mathcal{X}} L_{UV}^T [\Sigma^{-1} \otimes (f(x)f^T(x))] L_{UV} d\xi(x) \\ &= L_{UV}^T [\Sigma^{-1} \otimes M_f(\xi)] L_{UV}, \end{aligned}$$

where

$$M_f(\xi) = \int_{\mathcal{X}} f(x)f^T(x)d\xi(x) \tag{2.10}$$

is the information matrix of $\xi$ under the following single-response linear model with homoscedastic errors:

$$y(x) = f^T(x)\beta + e. \tag{2.11}$$

The above reformulations are demonstrated by the three examples given in the previous section. The linear and quadratic model (2.2) can be represented in the form of (2.8), with

$$f(x) = (1, x, x^2)^T,\ U_1 = (I_2, 0_{2\times 1}),\ U_2 = I_3,\ V_1 = (I_2, 0_{2\times 3}),\ V_2 = (0_{3\times 2}, I_3).$$

The Berman model in (2.3) can be represented in the form of (2.8), with

$$f(t) = (1, \cos t, \sin t)^T,\ U_1 = diag(1, 1, -1),\ U_2 = (e_1, e_3, e_2),$$
$$V_1 = (e_1, 0_{3\times 1}, e_2, e_3),\ V_2 = (0_{3\times 1}, I_3),$$

where $e_i$ is the $i$-th unit vector in $\mathbb{R}^3$, that is, $e_i$ has all-zero elements, except the $i$-th, which is unity. The parallel linear model in (2.4) can be represented in the form of (2.8), with

$$f(x) = (1, x_1, x_2)^T, \quad U_1 = (I_2, 0_{2\times 1}),\ U_2 = (e_1, 0_{2\times 1}, e_2),$$
$$V_1 = U_2, \qquad V_2 = (0_{2\times 1}, I_2),$$

where $e_i$ is the $i$-th unit vector in $\mathbb{R}^2$.

## 3. Admissible Designs

To discuss the admissibility of designs in multiresponse linear models, we start with the concept of admissibility introduced by Pukelsheim (1993). We first introduce some notation: $A \geq 0$ means that $A$ is a positive semidefinite (i.e., nonnegative definite) matrix. Two matrices $A, B$ are said to satisfy the inequality $A \geq B$ in the Loewner partial ordering if $A - B$ is positive semidefinite. Moreover, we use NND($p$) for the set of nonnegative definite matrices of order $p$, and Sym($p$) for the set of symmetric matrices of order $p$.

**Definition 1.** *An information matrix $M \in \mathcal{M}(\Xi)$ is called admissible in $\mathcal{M}(\Xi)$ when every competing information matrix $A \in \mathcal{M}(\Xi)$, with $A \geq M$, is actually equal to $M$. A design $\xi$ is called admissible in $\Xi$ when its information matrix $M(\xi)$ is admissible in $\mathcal{M}(\Xi)$.*

**Definition 2.** *A criterion function $\phi$ on NND(s) is a function $\phi$ : NND(s)*

$\to \mathbb{R}$ *that is positively homogeneous, superadditive, nonnegative, nonconstant, and upper semicontinuous.*

The following two lemmas provide the basic tools for the main results on the admissibility of designs in multiresponse linear models.

**Lemma 1.** *Let* $A$ *be an* $n \times n$ *positive semidefinite matrix and* $C$ *be a* $p \times n$ *matrix of rank* $q$ ($q \le p$)*. Then,*

a. $CAC^T \geq 0$*; in particular,* $CAC^T \gneq 0$ *if* $A \gneq 0$ *and* $C$ *is full column rank.*

b. $CA = 0$ *if* $CAC^T = 0$.

**Lemma 2.** *Suppose* $k \le p$*. Let* $T$ *be an* $r \times r$ *positive definite matrix,* $L$ *an* $rk \times p$ *matrix of rank* $p$*, and* $e_i$ *the* $i$*-th unit vector in* $\mathbb{R}^p$*. If* $e_{sk+1}, \ldots, e_{(s+1)k} \in$ *Ran(L) for some* $s \geq 0$*, then*

$$A \gneq B \Longleftrightarrow L^T[T \otimes A]L \gneq L^T[T \otimes B]L.$$

*Proof.* By part (a) of Lemma 1, if $A \gneq B$, then $L^T[T \otimes A]L \gneq L^T[T \otimes B]L$.

To prove the converse of the statement, we suppose that $L$ is partitioned as $[L_1^T, \ldots, L_r^T]^T$, where $L_i$ is a $k \times p$ matrix. If $e_{sk+1}, \ldots, e_{(s+1)k} \in \mathrm{Ran}(L)$, for some $s \geq 0$, then there exist $c_j \neq 0$ $(j = 1, \ldots, k)$ in $\mathbb{R}^p$, such that $Lc_j = e_{sk+j}$. Thus, $L_{s+1}C = I_k$ and $L_iC = 0$ $(i \neq s+1)$, where $C = [c_1, \ldots, c_k]$. If $L^T[T \otimes A]L \gneq L^T[T \otimes B]L$, then $C^TL^T[T \otimes A]LC \gneq C^TL^T[T \otimes B]LC$, that is, $T_{s+1,s+1}A \gneq T_{s+1,s+1}B$. Hence, $A \gneq B$ because $T_{s+1,s+1}$ is the $(s+1)$-th diagonal element of matrix $T$, which is positive definite. The proof is complete.

The first result on admissibility in the set $\Xi$ of all designs is about the location of the support points of admissible designs. To this end, we define the Elfving set by

$$\mathcal{R}_f = \mathrm{conv}\left(\{f(x) \mid x \in \mathcal{X}\} \cup \{-f(x) \mid x \in \mathcal{X}\}\right), \tag{3.1}$$

where conv($c$) denotes the convex hull of vectors $c \in \mathbb{R}^k$. The Elfving set $\mathcal{R}_f$ is a symmetric compact convex subset of $\mathbb{R}^k$ that contains the origin in its relative interior.

The following theorem states that in order to find optimal support points, we need to search the "extreme points" of the Elfving set $\mathcal{R}_f$ only.

**Theorem 1.** *Let* $\widetilde{\mathcal{R}}_f$ *be the set consisting of extreme points of the Elfving set* $\mathcal{R}_f$ *that do not lie on a straight line connecting any other two distinct points of the Elfving set* $\mathcal{R}_f$*. Then, for every design* $\eta \in \Xi$ *with support not included in*

$\widetilde{\mathcal{R}}_f$, *there exists a design* $\xi \in \Xi$ *with support included in* $\widetilde{\mathcal{R}}_f$, *such that*

$$M(\xi) \underset{\neq}{\geq} M(\eta). \tag{3.2}$$

*Proof.* From Theorem 8.5 in Pukelsheim (1993), there is a design $\xi \in \Xi$ such that

$$M_f(\xi) \underset{\neq}{\geq} M_f(\eta). \tag{3.3}$$

Hence,

$$\Sigma^{-1} \otimes M_f(\xi) \underset{\neq}{\geq} \Sigma^{-1} \otimes M_f(\eta),$$

and then

$$M(\xi) = L_{UV}^T[\Sigma^{-1} \otimes M_f(\xi)]L_{UV} \underset{\neq}{\geq} L_{UV}^T[\Sigma^{-1} \otimes M_f(\eta)]L_{UV} = M(\eta),$$

by part (a) of Lemma 1.

**Theorem 2.** *Let* $\phi$ *be a criterion function. If there is a* $\phi$*-optimal information matrix* $M_f$ *for the* $k$*-dimensional full parameter vector* $\beta$ *under the single-response model* (2.11) *in* $\mathcal{M}_f(\Xi)$*, then there exists a* $\phi$*-optimal design* $\xi$ *for* $\theta$ *under the multiresponse model* (2.1) *in* $\Xi$ *such that its support size,* $\sharp$ *supp*$(\xi)$*, is bounded according to*

$$\frac{p}{r} \leq \sharp\, supp(\xi) \leq \min\left(\frac{k(k+1)}{2}, \frac{p(p+1)}{2}\right). \tag{3.4}$$

*Proof.* By Corollary 8.3 in Pukelsheim (1993), for a design $\eta$ that has $M_f$ as its information matrix, there is an improved design $\xi$ with support size bounded from above by $k(k+1)/2$ and some $\delta \geq 1$ such that

$$\begin{aligned} \phi(M(\xi)) &= \phi(L_{UV}^T[\Sigma^{-1} \otimes M_f(\xi)]L_{UV}) \\ &= \phi(L_{UV}^T[\Sigma^{-1} \otimes (\delta M_f(\eta))]L_{UV}) \\ &= \delta\phi(L_{UV}^T[\Sigma^{-1} \otimes M_f(\eta)]L_{UV}) \\ &= \delta\phi(M(\eta)) \\ &\geq \phi(M(\eta)). \end{aligned} \tag{3.5}$$

On the other hand, Theorem 5.1.1 in Fedorov (1972) yields that

$$\frac{p}{r} \leq \sharp\, \mathrm{supp}(\xi) \leq \frac{p(p+1)}{2}. \tag{3.6}$$

This completes the proof of the theorem.

**Theorem 3.** *Suppose* $k \leq p$*. If* $e_{sk+1}, \ldots, e_{(s+1)k} \in Ran(L_{UV})$*, for some* $s \geq 0$*, then a design* $\xi$ *is admissible in* $\Xi$ *for the multiresponse model* (2.1) *if and only* *if $\xi$ is admissible in $\Xi$ for the single-response model* (2.11).

*Proof.* The proof follows immediately from Lemma 2 and Definition 1.

To illustrate the results given by Theorems 1–3, let us consider the linear and quadratic model in Example 1. The Elfving set corresponding to model (2.2) is given by

$$\mathcal{R}_1 = \mathrm{conv}\left(\{f(x) \mid x \in [-1,1]\} \cup \{-f(x) \mid x \in [-1,1]\}\right), \tag{3.7}$$

where $f(x) = (1, x, x^2)^T$. Its graph is shown in Pukelsheim (1993). In order to find optimal designs, we need only consider designs supported on the extreme points of $\mathcal{R}_1$. Furthermore, the support size is not more than six by Theorem 2. Finally, note that $e_{sk+1}, \ldots, e_{(s+1)k} \in \mathrm{Ran}(L_{UV})$ for $s = 1$. Then the admissible designs in the linear and quadratic model (2.2) are described by the following corollary.

**Corollary 1.** *A design $\xi \in \Xi$ is admissible in the linear and quadratic model* (2.2) *on the experimental region* $[-1, 1]$ *if and only if $\xi$ has at most one support in the open interval* $(-1, 1)$.

*Proof.* The proof follows immediately from Theorem 3 above and Claim 10.7 in Pukelsheim (1993).

## 4. Invariant Designs

In order to discuss the invariance of designs for the multiresponse model, we need the following concepts.

**Definition 3.** *The design problem for $\theta$ in $\mathcal{M}(\Xi)$ is said to be $\mathcal{Q}$-invariant when $\mathcal{Q}$ is a subgroup of the general linear group of order $p$, GL(p), and all transformations $Q \in \mathcal{Q}$ fulfill*

$$Q\mathcal{M}(\Xi)Q^T = \mathcal{M}(\Xi). \tag{4.1}$$

**Definition 4.** *A criterion function $\phi$ on NND(p) is called $\mathcal{H}$-invariant when $\mathcal{H}$ is a subgroup of the general linear group GL(p) and all transformations $H \in \mathcal{H}$ fulfill*

$$\phi(C) = \phi(HCH^T) \ \text{ for all } \ C \in NND(p). \tag{4.2}$$

**Definition 5.** *Let L: NND(s) $\to$ Sym(p) be the mapping $L(B) = L^T BL$, where $L$ has full column rank $p$. Assume $\mathcal{Q}$ is a subgroup of the general linear group*

*GL(s), and that there exists a group homomorphism $H$ from $\mathcal{Q}$ into GL(p), such that*

$$L(QBQ^T) = H(Q)L(B)H(Q)^T, \quad \text{for all} \;\; B \in NND(s), Q \in \mathcal{Q}, \tag{4.3}$$

*holds for the matrix $H(Q)$ in the image group $\mathcal{H}_{\mathcal{Q}} = \{H(Q)|Q \in \mathcal{Q}\}$. Then, the mapping $L$ is said to be $\mathcal{Q} - \mathcal{H}_{\mathcal{Q}}$-equivariant.*

To present the main result on the invariance of designs for multiresponse linear models, we need the following lemma.

**Lemma 3.** *Let $L_T$ : NND(k) $\to$ Sym(p) be the matrix mapping $L_T(A) = L^T(T \otimes A)L$ corresponding to an $rk \times p$ matrix $L$ of full column rank $p$ and a positive-definite matrix $T$ of order $r$. Assume $\mathcal{Q}$ is a subgroup of the general linear group GL(k). Define $N_Q$ : NND(rk) $\to$ NND(rk) by $N_Q(B) = (I_r \otimes Q)B$. Denote by $\mathcal{N}_{\mathcal{Q}}$ the set $\{N_Q|Q \in \mathcal{Q}\}$. We then have the following claims.*

a. *(Equivariance) There exists a group homomorphism $H : \mathcal{Q} \to$ GL(p) such that $L_T$ is equivariant under $H$,*

$$L_T(QAQ^T) = H(Q)L_T(A)H(Q)^T, \quad \text{for all} \;\; A \in NND(k), Q \in \mathcal{Q}, \tag{4.4}$$

*if the range of $L$ is invariant under each transformation $N_Q \in \mathcal{N}_{\mathcal{Q}}$,*

$$Ran(N_Q^T L) = Ran(L) \;\; \text{for all} \;\; N_Q \in \mathcal{N}_{\mathcal{Q}}. \tag{4.5}$$

b. *(Uniqueness) Suppose $L_T$ is equivariant under the group homomorphism $H : \mathcal{Q} \to$ GL(p). Then, $H(Q)$ or $-H(Q)$ is the unique nonsingular $p \times p$ matrix $H$ that satisfies $N_Q^T L = LH$, for all $N_Q \in \mathcal{N}_{\mathcal{Q}}$.*

c. *(Orthogonal transformation) Suppose $L_T$ is equivariant under the group homomorphism $H : \mathcal{Q} \to$ GL(p). If matrix $L$ fulfills $L^T L = I_p$ and $Q \in \mathcal{Q}$ is an orthogonal matrix of order $k$, then $H(Q) = \pm L^T N_Q^T L$ is an orthogonal matrix of order $p$.*

*Proof.* We only prove part (a). Let $g_T$ : NND($k$) $\to$ NND($rk$) be the matrix mapping $g_T(A) = T \otimes A$. Define $N$ :$\mathcal{Q} \to \mathcal{N}_{\mathcal{Q}}$ by $N(Q) = N_Q$. Then, $N$ is a group isomorphism such that the mapping $g_T$ is $\mathcal{Q} - \mathcal{N}_{\mathcal{Q}}$-equivariant under $N$; that is,

$$g_T(QAQ^T) = N(Q)g_T(A)N(Q)^T, \quad \text{for all} \;\; A \in \text{NND}(k), Q \in \mathcal{Q}. \tag{4.6}$$

On the other hand, by a similar argument to Lemma 13.5 in Pukelsheim (1993), there exists a group homomorphism $\overline{H} : \mathcal{N}_{\mathcal{Q}} \to$ GL($p$) such that the mapping $L$

is equivariant under $\overline{H}$,

$$L(N_Q B N_Q^T) = \overline{H}(N_Q)L(B)\overline{H}(N_Q)^T, \quad \text{for all} \;\; B \in \text{NND}(rk), \; N_Q \in \mathcal{N}_{\mathcal{Q}}, \tag{4.7}$$

if and only if the range of $L$ is invariant under each transformation $N_Q \in \mathcal{N}_{\mathcal{Q}}$,

$$\text{Ran}(N_Q^T L) = \text{Ran}(L) \;\; \text{for all} \;\; N_Q \in \mathcal{N}_{\mathcal{Q}}. \tag{4.8}$$

Therefore, $H = \overline{H} \circ N$: $\mathcal{Q} \to \text{GL}(p)$ is a group homomorphism such that the mapping $L_T$ is equivariant under $H$.

The set

$$\mathcal{H}_{\mathcal{Q}} = \left\{ H \in \text{GL}(p) \mid N_Q^T L = LH \;\; \text{for some} \;\; N_Q \in \mathcal{N}_{\mathcal{Q}} \right\} \tag{4.9}$$

is called the equivariance group induced by the $\mathcal{N}_{\mathcal{Q}}$-invariance of the design problem for the multiresponse model (2.1) in $\mathcal{M}(\Xi)$. Accordingly, the mapping $L_T$ is then said to be $\mathcal{Q} - \mathcal{H}_{\mathcal{Q}}$-equivariant.

The main result on invariant designs for the multiresponse model (2.1) is an immediate consequence of Lemma 3.

**Theorem 4.** *Let $\mathcal{Q}$ be a subgroup of the general linear group GL(k) and $\mathcal{N}_{\mathcal{Q}}$ the set $\{N_Q | Q \in \mathcal{Q}\}$. If all transformations $Q \in \mathcal{Q}$ fulfill*

$$Q\mathcal{M}_f(\Xi)Q^T = \mathcal{M}_f(\Xi) \tag{4.10}$$

*and*

$$Ran(N_Q^T L_{UV}) = Ran(L_{UV}) \;\; \textit{for all} \;\; N_Q \in \mathcal{N}_{\mathcal{Q}}, \tag{4.11}$$

*then the design problem for the multiresponse model* (2.1) *in* $\mathcal{M}(\Xi)$ *is* $\mathcal{H}_{\mathcal{Q}}$-*invariant.*

As an example, consider the linear and quadratic model (2.2) and the reflection transformation acting on the experimental domain $\mathcal{X} : x \to R(x) = -x$. The reflection $R(x) = -x$ leads to the $3 \times 3$ matrix $Q_R$=diag $(1, -1, 1)$, which reverses the sign of the linear component $x$ in the power vector $f(x) = (1, x, x^2)^T$; that is, $f(-x) = (1, -x, x^2)^T = Q_R f(x)$. Then, the reflection $R(x) = -x$ and the identity transformation induce a group $\mathcal{Q}$ of order 2:

$$\mathcal{Q} = \{I_3, Q_R\} = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \;\; \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\} \subset GL(3). \tag{4.12}$$

Because $L_{UV} = (e_1, e_2, e_4, e_5, e_6)$ and $(I_2 \otimes Q_R)L_{UV} = (e_1, -e_2, e_4, -e_5, e_6)$, where $e_i$ is the $i$-th unit vector in $\mathbb{R}^6$, this implies that $L_{UV}$ and $(I_2 \otimes Q_R)L_{UV}$ have the same range. By Theorem 4, this means that the design problem for the linear and quadratic model (2.2) is $\mathcal{H}_{\mathcal{Q}}$-invariant. Here, the equivariant group $\mathcal{H}_{\mathcal{Q}}$ is of order 2, as is $\mathcal{Q}$, containing the identity $I_5$ and $H_R = \text{diag }(1, -1, 1, -1, 1)$, which is obtained by part (c) of Lemma 3, because $L_{UV}^T L_{UV} = I_5$ and $Q_R \in \mathcal{Q}$ is an orthogonal matrix.

Let $\mathcal{H}$ be a finite subgroup of GL$(s)$ of order $\sharp\mathcal{H}$ and $\overline{C} : \text{Sym}(s) \to \text{Sym}(s)$ be the balancing operator defined by

$$\overline{C} = \frac{1}{\sharp\mathcal{H}} \sum_{H \in \mathcal{H}} HCH^T \qquad \text{for all} \quad C \in NND(s). \tag{4.13}$$

If $\phi$ is an $\mathcal{H}$-invariant criterion function, then the balancing operator leads to an improvement of a given information matric $C$,

$$\phi(\overline{C}) = \phi\left(\frac{1}{\sharp\mathcal{H}} \sum_{H \in \mathcal{H}} HCH^T\right) \geq \frac{1}{\sharp\mathcal{H}} \sum_{H \in \mathcal{H}} \phi(HCH^T) = \phi(C), \tag{4.14}$$

utilizing the concavity and $\mathcal{H}$-invariance of the criterion function $\phi$. Together with Corollary 1, we obtain a complete class $\Xi_{com}$ with minimum support size for the linear and quadratic model (2.2), composed of the following designs:

$$\xi = \begin{Bmatrix} -1 & 0 & 1 \\ w & 1-2w & w \end{Bmatrix}, \quad w \in \left[0, \frac{1}{2}\right]. \tag{4.15}$$

Note that the equality condition (4.11) is only sufficient. When it does not hold, there may exist a group $\mathcal{H}_{\mathcal{Q}}$ such that the design problem for the multiresponse model (2.1) in $\mathcal{M}(\Xi)$ is $\mathcal{H}_{\mathcal{Q}}$-invariant. A simple example is provided by the Berman model in (2.3), where the group $\mathcal{Q}$, induced by the reflection $R(t) = -t$ and the identity transformation, has order 2:

$$\mathcal{Q} = \{I_3, Q_R\} = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \right\} \subset GL(3). \tag{4.16}$$

Because $L_{UV} = (e_1, e_4, e_2 + e_6, -e_3 + e_5)$ and $(I_2 \otimes Q_R)L_{UV} = (e_1, e_4, e_2 - e_6, e_3 + e_5)$, where $e_i$ is the $i$-th unit vector in $\mathbb{R}^6$, it is easy to see that $\text{Ran}(L_{UV})$ is not the same as $\text{Ran}((I_2 \otimes Q_R)L_{UV})$. However, there exists a group $\mathcal{H}$ of order 2:

$$\mathcal{H} = \{I_4, H_R\} = \left\{ \begin{pmatrix} I_2 & 0 \\ 0 & I_2 \end{pmatrix}, \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix} \right\} \subset GL(4). \tag{4.17}$$

To this end, for every design $\xi \in \Xi$, we consider the *reflected design* $\xi^R$ given by $\xi^R(t) = \xi(-t)$, for all $t \in \mathcal{X} = [-\alpha/2, \alpha/2]$. The information matrix of $\xi$ is

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

where

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

The information matrix of $\xi^R$ is

$$\begin{aligned} M(\xi^R) &= \begin{pmatrix} I_2 & A(\xi^R) \\ A^T(\xi^R) & I_2 \end{pmatrix} = \begin{pmatrix} I_2 & A^T(\xi) \\ A(\xi) & I_2 \end{pmatrix} \\ &= \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix} \begin{pmatrix} I_2 & A(\xi) \\ A^T(\xi) & I_2 \end{pmatrix} \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix}^T. \end{aligned}$$

That is, the information matrix $M(\xi^R)$ is obtained from $M(\xi)$ by the congruence transformation,

$$M(\xi^R) = H_R M(\xi) H_R^T.$$

Consequently, $\mathcal{M}(\Xi)$ is invariant under transformation $H_R$. According to Theorem 3 and Lemma 1 in Wu (2002), we obtain a complete class $\Xi_{com}$ with minimum support size for the Berman model (2.3), composed of the following designs:

$$\xi = \begin{Bmatrix} -t & 0 & t \\ w & 1-2w & w \end{Bmatrix}, \quad t \in \left[0, \frac{\alpha}{2}\right], \quad w \in \left[0, \frac{1}{2}\right]. \tag{4.18}$$

## 5. Elfving's Theorem for $D$-optimality

In this section, we establish Elfving's theorem for $D$-optimality for multiresponse linear models. Define an Elfving set for the multiresponse linear model (2.1) by

$$\mathcal{R}_p = \operatorname{conv} \left\{ F^T(x) \Sigma^{-1/2} K \;\middle|\; x \in \mathcal{X}, K \in \mathbb{R}^{r \times p}, \|K\| = 1 \right\} \subseteq \mathbb{R}^{p \times p}, \tag{5.1}$$

where $\operatorname{conv}(B)$ denotes the convex hull of matrices $B \subseteq \mathbb{R}^{p\times p}$, and $\|K\|$ is the Frobenius norm of the matrix $K$; that is, $\|K\|^2 = \operatorname{tr}(K^T K)$. Note that the Elfving set $\mathcal{R}_p$ is a compact, symmetric, and convex subset of $\mathbb{R}^{p\times p}$ and contains the origin 0.

**Theorem 5.** *A design $\xi = \{x_v, w_v\}_{v=1}^s$ is $D$-optimal in $\Xi$ for the multiresponse linear model* (2.1) *if and only if $(pM(\xi))^{-1/2} \in \mathbb{R}^{p\times p}$ is a supporting hyperplane of the Elfving set $\mathcal{R}_p$, with supporting points $F^T(x_v)\Sigma^{-1/2}K_v$, where $K_v = (p\Sigma)^{-1/2} F(x_v) M^{-1/2}(\xi)$, for $v = 1, \ldots, s$.*

The proof Theorem 5 is based on the general equivalence theorem for $D$-optimality in multiresponse linear models (see Theorem 5.2.1 in Fedorov (1972)) and is similar to that of Theorem 3 in Liu, Yue and Lin (2013). Thus, the details of the proof are omitted here. To illustrate Theorem 5, we consider the parallel linear model given in (2.4).

The $D$-optimal design for estimating $\theta$ in model (2.4) is

$$\xi_D^* = \left\{ \begin{matrix} (-1,1) & (1,-1) \\ 1/2 & 1/2 \end{matrix} \right\} \quad \text{if } \rho > 0, \tag{5.2}$$

and

$$\xi_D^* = \left\{ \begin{matrix} (-1,-1) & (1,1) \\ 1/2 & 1/2 \end{matrix} \right\} \quad \text{if } \rho < 0 \tag{5.3}$$

(see Huang et al. (2006)). Now, Theorem 5 provides another way to verify that $\xi_D^*$ is $D$-optimal for model (2.4). Only the case $\rho > 0$ is shown below. The case $\rho < 0$ can be treated in a similar way. The information matrix of design $\xi_D^*$ is

$$M(\xi_D^*) = \frac{1}{1-\rho^2} \begin{pmatrix} 1 & -\rho & 0 \\ -\rho & 1 & 0 \\ 0 & 0 & 2(1+\rho) \end{pmatrix}. \tag{5.4}$$

Let $H(x) = (H_{ij})_{3\times 3} = F^T(x)\Sigma^{-1/2}K$, for $x = (x_1, x_2) \in \mathcal{X} = [-1,1]^2$ and $K \in \mathbb{R}^{2\times 3}$, with $\|K\| = 1$. The Elfving set $\mathcal{R}_3$ defined in (5.1) is given by Liu, Yue and Lin (2013) as follows:

$$\mathcal{R}_3 = \left\{ (H_{ij})_{3\times 3} \;\middle|\; \sum_{i=1}^{2}\sum_{j=1}^{3} H_{ij}^2 + 2\rho \sum_{j=1}^{3} H_{1j}H_{2j} \le 1, \right.$$
$$\left. |H_{3j}| \le |H_{1j} - H_{2j}|,\ j = 1,2,3 \right\},$$

and the boundary of $\mathcal{R}_3$ is obtained from the points $x = (-1,1)$ and $x = (1,-1)$. Therefore, $(-1,1)$ and $(1,-1)$ are the support points of the $D$-optimal design in the case $\rho > 0$. Corresponding to the two support points $(-1,1)$ and $(1,-1)$, we take

$$K_1 = (p\Sigma)^{-1/2}F(-1,1)M^{-1/2}(\xi_D^*) \quad \text{and} \quad K_2 = (p\Sigma)^{-1/2}F(1,-1)M^{-1/2}(\xi_D^*).$$

From (5.4) and the Cauchy–Schwarz inequality, we get

$$(\mathrm{tr}\{(pM(\xi_D^*))^{-1/2}F^T(x)\Sigma^{-1/2}K\})^2$$
$$\le \mathrm{tr}(K^TK)\mathrm{tr}\{(pM(\xi_D^*))^{-1}F^T(x)\Sigma^{-1}F(x)\}$$

$$= \frac{4 + 4\rho + x_1^2 + x_2^2 - 2\rho x_1 x_2}{6 + 6\rho} \leq 1,$$

for all $x \in \mathcal{X}$, whenever the matrix $K$ satisfies the equation $\|K\| = 1$. Moreover, a straightforward calculation gives

$$\mathrm{tr}\{(pM(\xi_D^*))^{-1/2} F^T(-1,1)\Sigma^{-1/2} K_1\} = 1,$$

$$\mathrm{tr}\{(pM(\xi_D^*))^{-1/2} F^T(1,-1)\Sigma^{-1/2} K_2\} = 1,$$

and $\|K_v\|^2 = \mathrm{tr}(K_v^T K_v) = 1$, for $v = 1, 2$. It follows that $\mathrm{tr}(M^{-1/2}(\xi_D^*)H) \leq 1$ for every $H \in \mathcal{R}_3$, and that each $K_v$ is on the boundary of $\mathcal{R}_3$. Thus, by Theorem 5, $\xi_D^*$ is $D$-optimal for model (2.4).

## 6. Concluding Remarks

Motivated by real applications, it is increasingly recognized that the multiresponse model is a useful tool for analyzing data from experiments with a multiple-outcome variable. While numerous excellent studies examine admissibility and invariance in single-response models, few address multiresponse models at the design stage.

In this study, we obtained the necessary and sufficient conditions for a design to be admissible and invariant, which are always helpful when investigating optimal designs for multiresponse linear models. We also established Elfving's theorem for $D$-optimality, which can be used to characterize $D$-optimal designs.

## Acknowledgments

This work was supported by NSFC grant (11471216, 11871143).

## References

Atkinson, A. C. and Bogacka, B. (2002). Compound and other optimum designs for systems of nonlinear differential equations arising in chemical kinetics. *Chemometr. Intell. Lab. Syst.* **61**, 17–33.

Berman, M. (1983). Estimating the parameters of a circle when angular differences are known. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **32**, 1–6.

Dette, H. (1993). Elfving's theorem for D-optimality. *Ann. Statist.* **21**, 753–766.

Dette, H. and Studden, W. J. (1993). Geometry of E-optimality. *Ann. Statist.* **21**, 416–433.

Dette, H. and Holland-Letz, T. (2009). A geometric characterization of c-optimal designs for heteroscedastic regression. *Ann. Statist.* **37**, 4088–4103.

Dette, H. and Melas, V. B. (2011). A note on the de la Garza phenomenon for locally optimal designs. *Ann. Statist.* **39**, 1266–1281.

Dette, H. and Schorning, K. (2013). Complete classes of designs for nonlinear regrssion models and principal representations of moment spaces. *Ann. Statist.* **41**, 1260–1267.

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

Elfving G. (1952). Optimum allocation in linear regression theory. *Ann. Math. Statist.* **23**, 255–262.

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

Holland-Letz, T., Dette, H. and Pepelyshev, A. (2011). A geometric characterization of optimal designs for regression models with correlated observations. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73**, 239–252.

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. (1990). Analysis of multiresponse experiments: A review. In *Statistical Design and Analysis of Industrial Experiments* (Edited by S. Ghosh), 231–246. Dekker, New York.

Khuri, A. I. (1996). Multiresponse surface methodology. In *Handbook of Statistics* **13** (Edited by S. Ghosh and C. R. Rao).

Kiefer, J. and Wolfowitz, J. (1959). Optimum designs in regression problem. *Ann. Math. Statist.* **30**, 271–294.

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

Liu, X., Yue, R.-X. and Hickernell, F. J. (2011). Optimality criteria for multiresponse linear models based on predictive ellipsoids. *Statist. Sinica* **21**, 421–432.

Liu, X., Yue, R.-X. and Lin, Dennis K. J. (2013). Optimal design for prediction in multiresponse linear models based on rectangular confidence region. *J. Statist. Plan. Inference* **143**, 1954–1967.

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

Silvey, S. D. and Titterington, D. M. (1973). A geometric approach to optimum design theory. *Biometrika* **60**, 21–32.

Smith K. (1918). On the standard deviations of adjusted and interpolated values of an observed polynomial function and its constants and the guidance they give towards a proper choice of the distribution of observations. *Biometrika* **12**, 1–85.

Soo, Y. W. and Bates, D. M. (1996). Multiresponse spline regression. *Comput. Statist. Data Anal.* **22**, 619–631.

Uciéski, D. and Bogacka, B. (2005). T-optimum designs for discrimination between two multiresponse dynamic models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **67**, 3–18.

Wu, H. (2002). Optimal designs for first-order trigonometric regression on a partial circle. *Statist. Sinica* **12**, 917–930.

Yang, M. (2010). On the de la Garza phenomenon. *Ann. Statist.* **38**, 2499–2524.

Yang, M. and Stufken, J. (2009). Support points of locally optimal designs for nonlinear models with two parameters. *Ann. Statist.* **37**, 518–541.

Yang, M. and Stufken, J. (2012). Identifying locally optimal designs for nonlinear models: A simple extension with profound consequences. *Ann. Statist.* **40**, 1665–1685.

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

E-mail: liuxin@dhu.edu.cn

College of Mathematics and Science, Shanghai Normal University, Shanghai 200234, P. R. China.

E-mail: yue2@shnu.edu.cn

(Received July 2017; accepted February 2018)