Journal of Statistical Planning and Inference 143 (2013) 1954–1967

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Optimal design for prediction in multiresponse linear models based on rectangular confidence region ☆

Xin Liu [a], Rong-Xian Yue [b,*], Dennis K.J. Lin [c]

[a] College of Science, Donghua University, Shanghai 201600, China
[b] College of Mathematics and Science, Shanghai Normal University, and Scientific Computing Key Laboratory of Shanghai Universities, Shanghai 200234, China
[c] Department of Statistics, The Pennsylvania State University, University Park, PA 16802, USA

ARTICLE INFO

Article history:
Received 10 January 2013
Received in revised form
4 June 2013
Accepted 18 June 2013
Available online 12 July 2013

Keywords:
Elfving's theorem
General equivalence theorem
Multiresponse linear models
Optimal design
Rectangular confidence region

ABSTRACT

This paper proposes a class of optimal design criteria for prediction in linear regression models with r responses based on the volume of the rectangular confidence region, termed R_L^r-optimality. A general equivalence theorem and a geometrical characterization of the R_L^r-optimal design are established, which are used to obtain or verify the R_L^r-optimality. Several examples are given for illustration.

1. Introduction

Consider the design problem for prediction in multiresponse linear regression models of the form

$$Y(x) = F(x)\theta + \varepsilon, \tag{1}$$

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. The experimental design region is $\mathcal{X}$, which is a compact subset in the q-dimensional Euclidean space.

Multiresponse experiments are frequently encountered in applications, such as chemical engineering, food science, manufacturing, biological and medical studies, etc. The optimal design problem for multiresponse models has been discussed by many authors. For example, Draper and Hunter (1966) developed a criterion for parameter estimation of a multiresponse model. Fedorov (1972, Chapter 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

☆ This work was partially supported by NSFC Grants (11071168, 11101077, 11201062), Special Funds for Doctoral Authorities of Education Ministry (20103127110002), Innovation Program of Shanghai Municipal Education Commission (11zz116), E-Institutes of Shanghai Municipal Education Commission (E03004), Shanghai Leading Academic Discipline Project, the Fundamental Research Funds for the Central Universities.
* Corresponding author. Tel.: +86 21 64323611.
E-mail address: yue2@shnu.edu.cn (R.-X. Yue).

0378-3758/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jspi.2013.06.012

multiresponse experiments. Gueorguieva et al. (2006) considered the optimal design problem for multivariate response pharmacokinetic models. Jin and Yue (2008) studied the D- and A-optimal designs for mixture multiresponse experiments. Wang and Yue (2008) developed an equivalent theorem of Bayesian optimal design for multiresponse linear regression models. Atashgah and Seifi (2009) proposed a unified formulation for the multiresponse optimal design problem using semi-definite programming and generated D-, A- and E-optimal designs. Zhu et al. (2010) considered robust designs against outliers for multiresponse models. Sagnol (2011) computed the multiresponse optimal design problem by second-order cone programming. Liu et al. (2011) proposed a new class of optimum design criteria for the linear regression model with r responses based on the volume of the predictive ellipsoid for the vector of responses on a predetermined interval $\mathcal{Z}$. The I_L^r-optimality criterion reduces to I_L-optimality proposed by Dette and O'Brien (1999) in a single response situation.

In this paper, an alternative criterion for prediction is proposed in the multiresponse linear regression model (1) based on the Bonferroni confidence rectangle. This optimality is a generalization of R-optimality proposed by Dette (1997) to the case of multiresponse experiments.

The paper is organized as follows. Section 2 introduces the optimality criteria for linear regression models with r responses, termed R_L^r-optimality. An equivalence theorem and a new geometric characterization of Elfving (1952) type for the R_L^r-optimal design problem are derived in Section 3. Several illustrative examples are given in Section 4 and some concluding remarks are given in Section 5. All proofs are presented in the Appendix.

2. Development of R_L^r-optimality

Throughout the paper we consider approximate designs of the form

$$\xi = \left\{ \begin{matrix} x_1 & \cdots & x_s \\ w_1 & \cdots & w_s \end{matrix} \right\},$$

or $\xi = \{x_v, w_v\}_{v=1}^s$ for short, where the support points $x_1, \ldots, x_s$ are distinct elements of the design region $\mathcal{X} \subset \mathcal{R}^q$, and corresponding weights $w_1, \ldots, w_s$ are nonnegative real numbers which sum to unity. Denote the set of all approximate designs with positive semi-definite information matrix on $\mathcal{X}$ by Ξ. The information matrix of ξ for model (1) is

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

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

We assume the matrix $F(x)$ 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 the r predicted responses at z under the design ξ is

$$V(z,\xi) = F(z) M^{-1}(\xi) F^T(z). \tag{2}$$

Throughout the paper by $V_{ij}(z,\xi)$ we denote the (i,j)th entry of the $r \times r$ matrix $V(z,\xi)$, i.e.,

$$V_{ij}(z,\xi) = e_i^T V(z,\xi) e_j, \quad i,j \in \{1,2,\ldots,r\},$$

where e_i is the ith unit vector in $\mathcal{R}^r$. When there is no possibility of confusion we omit the dependence of $M(\xi)$, $V(z,\xi)$ and $F(z)$ on ξ and z.

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

$$\psi_L(\xi) = \left\{ \int_{\mathcal{Z}} \left(\prod_{i=1}^r V_{ii}(z,\xi) \right)^L d\mu(z) \right\}^{1/L} \tag{3}$$

over Ξ.

Remark 1. This definition can be extended to allow the case $L = \infty$ by taking $\psi_\infty(\xi) = \sup_{z \in \mathcal{Z}} \prod_{i=1}^r V_{ii}(z,\xi)$. It can be shown that $\psi_\infty(\xi) = \lim_{L\to\infty} \psi_L(\xi)$ if $\mathrm{supp}(\mu) = \mathcal{Z}$ and each element of the $r \times p$ matrix $F(z)$ of regression functions is continuous on $\mathcal{Z}$, where $\mathrm{supp}(\mu)$ denotes the set of support points of μ. Obviously, the R_∞^r-optimality criterion minimizes the maximum volume of the prediction rectangle. It is clear that the R_L^r-optimality is coincided with the I_L^r-optimality when $r=1$, and can be viewed as a generalization of G-optimality to multiresponse situations.

Remark 2. Yue and Liu (2010) show that I_L^r-optimal designs for hierarchically ordered system of regression models do not depend on the variance–covariance matrix of the responses. However, as will be shown in Section 4, the R_L^r-optimal designs may depend on the correlation of the responses.

Remark 3. Comparing with Kiefer's Φ_k class, a good property of R_L^r-optimality is that it is invariant with respect to model reparameterization. Thus the matrix $F(x)$ can be replaced by $G(x) := F(x)A$ for any nonsingular $p \times p$ matrix A and θ replaced by $\beta := A^{-1}\theta$. This is also noted for I_L-optimality by Dette and O'Brien (1999, Theorem 1).

3. An equivalence theorem and Elfving's theorem for R_L^r-optimality

The general equivalence theorem plays an important role in verifying optimality of a design. Here we present an equivalence theorem for R_L^r-optimality to characterize R_L^r-optimal designs.

Lemma 1. *Let $\mathcal{P}_s$ denote the set of all $n \times n$ positive definite matrices and A be a fixed $m \times n$ $(n \geq m)$ matrix. Then $g(B) := \prod_{i=1}^m (AB^{-1}A^T)_{ii} = \prod_{i=1}^m e_i^T AB^{-1}A^T e_i$ is convex on $\mathcal{P}_s$.*

This lemma is a special case of results in Gaffke and Heiligers (1996, p. 1153). From this lemma and (3), we have the following.

Lemma 2. *For the criterion function ψ_L defined by (3) we have*

(i) *ψ_L is convex on Ξ;*
(ii) *the directional derivative of ψ_L at ξ in the direction of $\overline{\xi}$, denoted $\Delta_{\psi_L}(\xi, \overline{\xi})$, is*
$$\Delta_{\psi_L}(\xi, \overline{\xi}) = r\psi_L(\xi) - \psi_L^{1-L}(\xi)\ \mathrm{tr}\{M^{-1}(\xi)M(\overline{\xi})M^{-1}(\xi)Q_L(\xi)\},$$
where
$$Q_L(\xi) = \int_{\mathcal{Z}} \left(\prod_{i=1}^r V_{ii}^L(z, \xi) \right) \sum_{i=1}^r \frac{F^T(z)e_i e_i^T F(z)}{V_{ii}(z, \xi)}\, d\mu(z); \tag{4}$$
(iii) *for any fixed ξ with nonsingular information matrix, the directional derivative*
$$\Delta_{\psi_L}(\xi, \overline{\xi}) = \int_{\mathcal{X}} \Delta_{\psi_L}(\xi, \delta_x)\, d\overline{\xi}(x), \quad (\xi, \overline{\xi}) \in \Xi \times \Xi$$
is linear in $\overline{\xi}$, where $\delta_x \in \Xi$ denotes the Dirac measure at x.

Note that ξ_L^* is R_L^r-optimal in Ξ if and only if $\inf_{x \in \mathcal{X}} \Delta_{\psi_L}(\xi_L^*, \delta_x) = 0$ (see Whittle, 1973). The following theorem can thus be established.

Theorem 1. *For all $L \in [1, \infty)$, a design $\xi_L^* \in \Xi$ is R_L^r-optimal in Ξ if and only if*
$$\sup_{x \in \mathcal{X}} \mathrm{tr}\{M^{-1}(\xi_L^*)F^T(x)\Sigma^{-1}F(x)M^{-1}(\xi_L^*)Q_L(\xi_L^*)\} = r\int_{\mathcal{Z}} \prod_{i=1}^r V_{ii}^L(z, \xi_L^*)\, d\mu(z). \tag{5}$$
Moreover, the supremum is achieved at the support points of ξ_L^*.

To compare R_L^r-optimal design to other designs, e.g., I_1^r- and I_∞^r-optimal designs obtained in Liu et al. (2011), we define the R_L^r-efficiency of a design ξ with respect to a R_L^r-optimal design ξ_L^* as
$$R_L^r - \mathrm{Eff}(\xi) = \frac{\psi_L(\xi_L^*)}{\psi_L(\xi)}. \tag{6}$$
It is clear that $0 \leq R_L^r - \mathrm{Eff}(\xi) \leq 1$ for all $\xi \in \Xi$. The following corollary provides a lower bound for $R_L^r - \mathrm{Eff}(\xi)$ that follows immediately from Theorem 1.

Corollary 1. *For $L \in [1, \infty)$, let*
$$\phi_L(x, \xi) = \frac{\mathrm{tr}\{M^{-1}(\xi)F^T(x)\Sigma^{-1}F(x)M^{-1}(\xi)Q_L(\xi)\}}{\int_{\mathcal{Z}} \prod_{i=1}^r V_{ii}^L(z, \xi)\, d\mu(z)}. \tag{7}$$
Then $R_L^r - \mathrm{Eff}(\xi) \geq 1 + r - \sup_{x \in \mathcal{X}} \phi_L(x, \xi)$.

In terms of the function $\phi_L(x, \xi)$ at (7), we can restate Theorem 1 as follows.

Theorem 1′. *For $L \in [1, \infty)$, a design $\xi_L^* \in \Xi$ is R_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^*.

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

$$\mathcal{A}(\xi)=\left\{z\in\mathcal{Z}\,\middle|\,\prod_{i=1}^{r}V_{ii}(z,\xi)=\sup_{z'\in\mathcal{Z}}\prod_{i=1}^{r}V_{ii}(z',\xi)\right\}.$$

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

$$\phi_\infty(x,\xi)=\mathrm{tr}\left\{M^{-1}(\xi)F^T(x)\Sigma^{-1}F(x)M^{-1}(\xi)\int_{\mathcal{A}(\xi)}\sum_{i=1}^{r}\frac{F^T(z)e_ie_i^TF(z)}{V_{ii}(z,\xi)}d\mu^*(z)\right\}. \tag{8}$$

Theorem 2. *A design $\xi_\infty^*\in\Xi$ is R_∞^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_∞^*.

Other useful methods for the determination of the optimal design are geometric characterizations. Elfving (1952) characterizes the c-optimal designs in single response experiments, and shows that the optimal design can be found at the intersection of a vectorial straight-line. This result was generalized by Studden (1971) to optimal designs for parameter systems $A^T\theta$. Sagnol (2011) gave a generalization to the case of multiresponse experiments. It is worth mentioning that Elfving-type characterizations are also available for other optimality criteria, and we refer the reader to, e.g., Dette (1993a,b, c, 1996), and Holland-Letz et al. (2009). We now present a geometric structure which can be used for the characterization of optimal designs with respect to the R_L^r-optimality criteria for multiresponse experiments.

Define an Elfving set by

$$\mathcal{R}_p=\mathrm{conv}\{F^T(x)\Sigma^{-1/2}K|x\in\mathcal{X},K\in\mathbb{R}^{r\times p},\|K\|=1\}\subseteq\mathbb{R}^{p\times p}, \tag{9}$$

where 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, i.e., $\|K\|^2=\mathrm{tr}(K^TK)$. $\mathcal{R}_p$ is a compact, symmetric and convex subset of $\mathbb{R}^{p\times p}$ and contains the point 0. Let $A(\xi)$ be an $p\times p$ matrix satisfying

$$A(\xi)A(\xi)^T=\begin{cases}\dfrac{1}{r}Q_L(\xi) & \text{for } L\in[1,+\infty),\\[2ex] \dfrac{1}{r}\int_{\mathcal{A}(\xi)}\sum_{i=1}^{r}\dfrac{F^T(z)e_ie_i^TF(z)}{V_{ii}(z,\xi)}d\mu^*(z) & \text{for } L=+\infty.\end{cases}$$

A geometric characterization of the R_L^r-optimal designs can be established.

Theorem 3. *For a given $L\in[1,\infty]$, a design $\xi=\{x_v,w_v\}_{v=1}^s\in\Xi$ is R_L^r-optimal in Ξ if and only if there exist a positive number γ and matrices $K_1,K_2,\ldots,K_s$ such that*

(i) $\gamma A(\xi)=\sum_{v=1}^{s}w_vF^T(x_v)\Sigma^{-1/2}K_v$;
(ii) *$\gamma A(\xi)$ is a boundary point of the set $\mathcal{R}_p$ with a supporting hyperplane $D\in\mathbb{R}^{p\times p}$*;
(iii) $\|K_v\|=1,\quad v=1,\ldots,s$.

4. Illustrative examples

In this section, we present three examples of R_1^r- and R_∞^r-optimal designs for bivariate (r=2) response models. Example 1 considers a linear and a quadratic regression model, and both R_1^r- and R_∞^r-optimal designs are constructed. In Example 2, optimal design for a multi-factor polynomial model is constructed and its optimality is proved by means of the equivalence theorem. Example 3 considers a parallel linear model with two responses, and shows that the D-optimal design is also R_1^r- and R_∞^r-optimal by Theorem 3. To simplify the notation in three examples, we replace R_L^r and I_L^r by R_L and I_L, respectively.

Example 1 (*Linear and quadratic regression*). Consider two responses and one controllable variable. The experimental region and the prediction region coincide, i.e., $\mathcal{X}=\mathcal{Z}=[-1,1]$, and the assumed regression model is

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

Denote Σ as the variance–covariance matrix of the random vector $\varepsilon=(\varepsilon_1,\varepsilon_2)^T$. Let Σ and its inverse Σ^{-1} be of the following forms:

$$\Sigma=\begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{pmatrix},\quad \Sigma^{-1}=\begin{pmatrix}\sigma^{11} & \sigma^{12}\\ \sigma^{21} & \sigma^{22}\end{pmatrix},$$

where $\sigma_{12}=\sigma_{21}=\rho\sqrt{\sigma_{11}\sigma_{22}}$. Further, let μ be the uniform measure on $\mathcal{Z}$, i.e., $d\mu(z)=\frac{1}{2}\,dz$.

For this model, the matrix $F(x)$ becomes

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

and the vector of parameters is $\theta=(\theta_{10},\theta_{11},\theta_{20},\theta_{21},\theta_{22})^T$. For a design ξ, the information matrix is then

$$M(\xi)=\begin{pmatrix}\sigma^{11}s_0 & \sigma^{11}s_1 & \sigma^{12}s_0 & \sigma^{12}s_1 & \sigma^{12}s_2\\ \sigma^{11}s_1 & \sigma^{11}s_2 & \sigma^{12}s_1 & \sigma^{12}s_2 & \sigma^{12}s_3\\ \sigma^{12}s_0 & \sigma^{12}s_1 & \sigma^{22}s_0 & \sigma^{22}s_1 & \sigma^{22}s_2\\ \sigma^{12}s_1 & \sigma^{12}s_2 & \sigma^{22}s_1 & \sigma^{22}s_2 & \sigma^{22}s_3\\ \sigma^{12}s_2 & \sigma^{12}s_3 & \sigma^{22}s_2 & \sigma^{22}s_3 & \sigma^{22}s_4\end{pmatrix},$$

where s_k is the kth moment of ξ, i.e., $s_k=\int_{\mathcal{X}}x^k\,d\xi$. Obviously we have $0\le s_2\le 1$ and $s_2^2\le s_4\le s_2$. Note that if $\tilde{\xi}$ denotes the reflection of ξ at the origin, then ξ and $\tilde{\xi}$ have the identical even moments, while a reversed sign for the odd moments. It follows that $M(\tilde{\xi})=TM(\xi)T$, with $T=\mathtt{Diag}(1,-1,1,-1,1)$. This implies that $M^{-1}(\tilde{\xi})=TM^{-1}(\xi)T$ and $\psi_L(\tilde{\xi})=\psi_L(\xi)$. Thus, if ξ is R_L-optimal, then $\tilde{\xi}$ is R_L-optimal. This implies that the symmetrized design $\xi^*=(\xi+\tilde{\xi})/2$ is alsoR_L-optimal due to the convexity of the R_L-criterion. It is thus sufficient to search for R_L-optimal designs among the symmetric designs on $\mathcal{X}$. For a symmetric design ξ we have $s_1=s_3=0$ and

$$h(x,\xi):=\prod_{i=1}^{2}V_{ii}(x,\xi)=\frac{\sigma_{11}\sigma_{22}(x^2+s_2)[(1-\rho^2)s_2x^4+(s_4+2\rho^2s_2^2-3s_2^2)x^2+s_2s_4-\rho^2s_2^3]}{s_2^2(s_4-s_2^2)}. \tag{11}$$

Noting that $\psi_L(\xi)=\frac{1}{2}(\int_{-1}^{1}h^L(x,\xi)\,dz)^{1/L}$ depends on ξ only through the moments s_2 and s_4, we denote $\psi_L(\xi)=\eta_L(s_2,s_4)$. Only the cases $L=1$ and $L=\infty$ will be considered.

1(a). For $L=1$, it is straightforward to show that

$$\eta_1(s_2,s_4)=\frac{\sigma_{11}\sigma_{22}[(105s_2^2+70s_2+21)s_4-105\rho^2s_2^4+35(\rho^2-3)s_2^3+21(\rho^2-2)s_2^2+15(1-\rho^2)s_2]}{105s_2^2(s_4-s_2^2)}$$

and

$$\frac{d\eta_1(s_2,s_4)}{ds_4}=\frac{\sigma_{11}\sigma_{22}(1-\rho^2)(-105s_2^3+35s_2^2+21s_2-15)}{105s_2(s_4-s_2^2)^2}<0.$$

It follows that for a fixed s_2 and $s_2^2\le s_4\le s_2$, $\eta_1(s_2,s_4)$ decreases in s_4. Consequently,

$$\eta_1(s_2,s_4)\ge\eta_1(s_2,s_2)=\frac{\sigma_{11}\sigma_{22}[-105\rho^2s_2^3+35\rho^2s_2^2+(21\rho^2+28)s_2+36-15\rho^2]}{105s_2^2(1-s_2)}$$

with equality holds if and only if $s_4=s_2$. Therefore, for the optimal moments, we have $s_4^*=s_2^*$ and $s_1=s_3=0$. However $s_4=s_2$ is only possible when the support points $\mathtt{supp}(\xi)\subset\{-1,0,1\}$ and since s_1 vanishes, -1 and 1 must appear equally often in ξ. Hence let s_2^* be a minimizer of $\eta_1(s_2,s_2)$ on $[0,1]$ then the following design

$$\xi_1^*=\begin{Bmatrix}-1 & 0 & 1\\ \frac{s_2^*}{2} & 1-s_2^* & \frac{s_2^*}{2}\end{Bmatrix} \tag{12}$$

is R_1-optimal. Note that the optimal moment s_2^* depends only on the correlation coefficient ρ, but not on σ_{11} and σ_{22}. In particular, if $\rho=0$ then $s_2^*=(2\sqrt{22}-5)/7$. For $0<|\rho|<1$, the value of s_2^* can be found numerically. Table 1 shows the values of s_2^* for various ρ in $(-1,1)$.

1(b). For $L=\infty$, it is not difficult to find that

$$\psi_\infty(\xi)=\max\{h(0,\xi),h(1,\xi)\}=\begin{cases}h(1,\xi) & (s_4\ge c(s_2));\\ h(0,\xi) & \text{otherwise,}\end{cases}$$

Table 1
Values of s_2^* in (12) for various ρ in $(-1,1)$.

$\lvert\rho\rvert$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.99
s_2^*	0.626	0.627	0.629	0.634	0.640	0.650	0.663	0.693	0.713	0.767	0.906

where $h(\cdot,\xi)$ is given in (11), and

$$c(s_2)=\frac{s_2[(\rho^2-1)+(2-\rho^2)s_2+(3-\rho^2)s_2^2]}{1+2s_2}.$$

Note that the minimum of $\psi_\infty(\xi)$ is attained when $h(0,\xi)=h(1,\xi)$, which simplifies to $s_4=c(s_2)$. Therefore, we have

$$\eta_\infty(s_2,s_4)\geq\eta_\infty(s_2,c(s_2))=1+\frac{(1-\rho^2)s_2^2}{c(s_2)-s_2^2}=\frac{3s_2^2+2s_2-1}{s_2^2+s_2-1},$$

and

$$\frac{d\eta_\infty(s_2,c(s_2))}{ds_2}=\frac{s_2^2-4s_2-1}{(s_2^2+s_2-1)^2}<0.$$

It follows that $\eta_\infty(s_2,c(s_2))$ decreases in s_2 on [0, 1], and thus

$$\frac{\rho^2-\sqrt{5\rho^4-20\rho^2+24}}{6-2\rho^2}\leq s_2\leq\frac{\rho^2+\sqrt{5\rho^4-20\rho^2+24}}{6-2\rho^2}$$

due to the restriction $s_2^2\leq s_4\leq s_2$. Thus, $\psi_\infty(\xi)=\eta_\infty(s_2,s_4)$ attains its minimum when $s_2=(\rho^2+\sqrt{5\rho^4-20\rho^2+24})/(6-2\rho^2)$ and $s_4=c(s_2)=s_2$. Therefore, R_∞-optimal design ξ_∞^* is as follows:

$$\xi_\infty^*=\left\{\begin{matrix}-1 & 0 & 1\\ \frac{s_2^*}{2} & 1-s_2^* & \frac{s_2^*}{2}\end{matrix}\right\}\quad\text{where } s_2^*=\frac{\rho^2+\sqrt{5\rho^4-20\rho^2+24}}{6-2\rho^2}. \tag{13}$$

Fig. 1 shows the optimal weights w^* at ±1 versus $|\rho|$ corresponding to the R_L-optimal design ξ_L^* for $L=1$ (solid line) and $L=\infty$ (dashed line), respectively.

It is worthwhile to compare the R_1- and R_∞-optimal designs with D-, I_1- and I_∞-optimal designs. The D-optimal design for model (10) on $\mathcal{X}=[-1,1]$ is (see Krafft and Schaefer, 1992)

$$\xi_D^*=\left\{\begin{matrix}-1 & 0 & 1\\ \frac{3}{8} & \frac{1}{4} & \frac{3}{8}\end{matrix}\right\}.$$

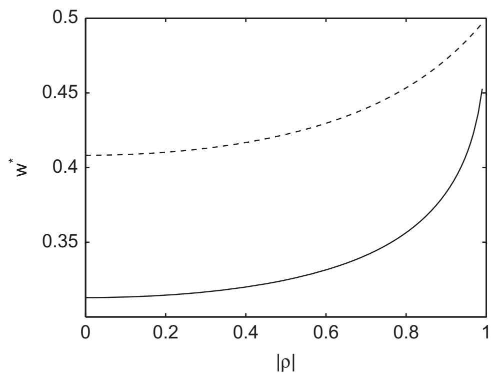

Fig. 1. The optimal weights w^* at ±1 versus $|\rho|$ corresponding to the R_L-optimal design ξ_L^* for $L=1$ (solid) and $L=\infty$ (dash), respectively, for model (10).

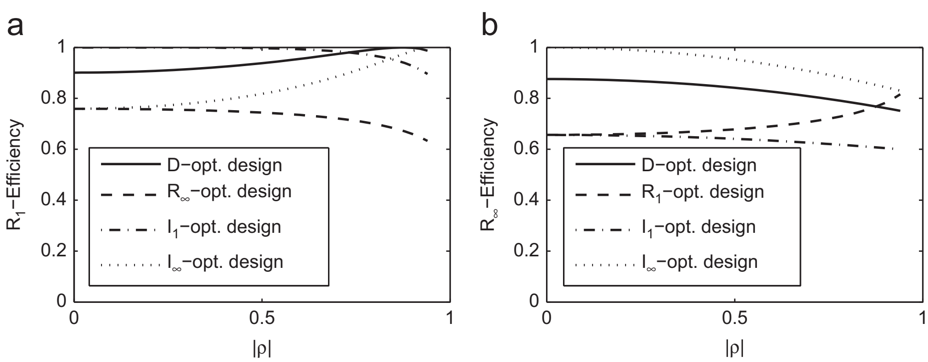

Fig. 2. (a) The R_1-efficiencies of ξ_D^* (solid), ξ_∞^* (dash), ζ_1^* (dash–dot) and ζ_∞^* (dot) with respect to ξ_1^*; (b) the R_∞-efficiencies of ξ_D^* (solid), ξ_1^* (dash), ζ_1^* (dash–dot) and ζ_∞^* (dot) with respect to ξ_∞^*.

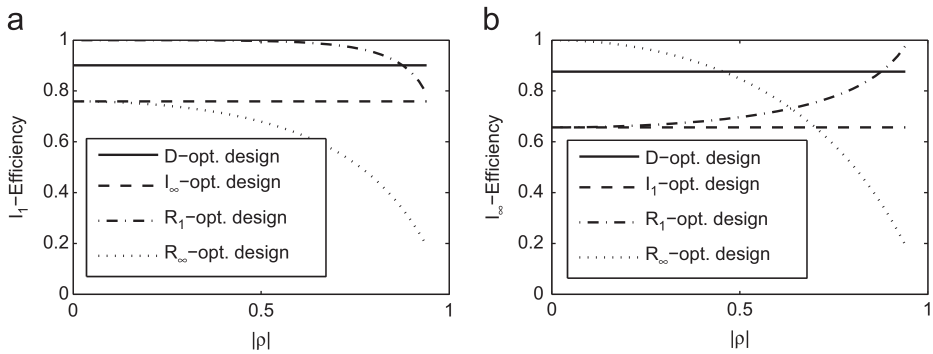

Fig. 3. (a) The I_1-efficiencies of ξ_D^* (solid), ζ_∞^* (dash), ξ_1^* (dash–dot) and ξ_∞^* (dot) with respect to ζ_1^*; (b) the I_∞-efficiencies of ξ_D^* (solid), ζ_1^* (dash), ξ_1^* (dash–dot) and ξ_∞^* (dot) with respect to ζ_∞^*.

The I_1- and I_∞-optimal designs for model (10) on $\mathcal{X}=\mathcal{Z}=[-1,1]$, denoted by ζ_1^* and ζ_∞^*, respectively, are as follows (see Liu et al., 2011):

$$\zeta_1^*=\left\{\begin{array}{ccc}-1 & 0 & 1\\ \frac{2\sqrt{22}-5}{14} & \frac{12-2\sqrt{22}}{7} & \frac{2\sqrt{22}-5}{14}\end{array}\right\},\quad \zeta_\infty^*=\left\{\begin{array}{ccc}-1 & 0 & 1\\ \frac{\sqrt{6}}{6} & 1-\frac{\sqrt{6}}{3} & \frac{\sqrt{6}}{6}\end{array}\right\}.$$

Note that the D-, I_1- and I_∞-optimal designs have the same structure as the R_1- and R_∞-optimal designs but do not depend on Σ. In addition, in the case $\rho=0$, the R_1- and R_∞-optimal designs coincide with the I_1- and I_∞-optimal designs, respectively. Compared with the D-optimal design, the R_∞-optimal design is always concentrated at the endpoints ± 1 since $s_2^*/2\geq 3/8$; while the R_1-optimal design puts more concentration at the origin when $|\rho|\leq 4\sqrt{3523}/271\approx 0.8761$.

Fig. 2 shows the R_1-efficiencies and R_∞-efficiencies of the D-, R_1-, R_∞-, I_1- and I_∞-optimal designs. These are calculated according to (6). Fig. 3 shows the I_1-efficiencies and I_∞-efficiencies of the D-, R_1-, R_∞-, I_1- and I_∞-optimal designs. The I_L-efficiency of a design ξ is the relative efficiency of ξ with respect to a I_L-optimal design under the I_L-optimality defined in Liu et al. (2011).

The following factors are observed for the multiresponse model (10):

(i) the R_L-optimal design for the multiresponse model depends on the covariance matrix Σ of the two responses only through the correlation coefficient ρ;
(ii) the loss of efficiency might be moderate or substantial, if a rectangular confidence region is constructed on the basis of a I_L^r-optimal design and a D-optimal design, or a confidence ellipsoid is constructed on the basis of a R_L^r-optimal design and a D-optimal design;
(iii) the performance of the I_∞-optimal design is worse than the D-, R_∞- and I_1-optimal designs for constructing rectangular confidence region based on R_1-optimality;
(iv) the performance of the I_1-optimal design is worse than the D-, R_1- and I_∞-optimal designs for constructing rectangular confidence region based on R_∞-optimality.

Example 2 (*Multi-factor polynomial*). Consider an experiment involving two responses with m controllable variables, and $\mathcal{X}=\mathcal{Z}=[-1,1]^m$. The assumed regression model is

$$\begin{cases} y_1(x)=\theta_{11}+\sum_{i=1}^m\theta_{1i}x_i+\varepsilon_1,\\ y_2(x)=\theta_{21}+\sum_{i=1}^m\theta_{2i}x_i+\sum_{1\leq i<j\leq m}\theta_{2ij}x_ix_j+\varepsilon_2.\end{cases}\tag{14}$$

Let the variance–covariance matrix of the two responses be Σ and its inverse be as in Example 1. The *uniform factorial*2^m-*design* ξ_U assigns equal weight 2^{-m} to each of the 2^m extreme points $(\pm 1,\pm 1,\ldots,\pm 1)$. It can be shown that ξ_U is R_L-optimal design for $L=1$ and $L=\infty$. We next verify its optimality by Theorem 1′ and Theorem 2.

Denote $g_1(x)=(1,x_1,\ldots,x_m)^T$ and $g_2(x)=(x_1x_2,\ldots,x_{m-1}x_m)^T$. Then

$$F(x)=\begin{pmatrix} g_1^T(x) & 0 & 0\\ 0 & g_1^T(x) & g_2^T(x)\end{pmatrix}.$$

The information matrix of ξ_U and its inverse are given by

$$M(\xi_U)=\begin{pmatrix}\Sigma^{-1}\otimes I_{m+1} & 0\\ 0 & \sigma^{22}I_{(m^2-m)/2}\end{pmatrix},\quad M^{-1}(\xi_U)=\begin{pmatrix}\Sigma\otimes I_{m+1} & 0\\ 0 & \frac{1}{\sigma^{22}}I_{(m^2-m)/2}\end{pmatrix},$$

where I_k denotes the $k\times k$ identity matrix. It follows that

$$V(z,\xi_U)=F(z)M^{-1}(\xi_U)F^T(z)=\begin{pmatrix}\sigma_{11}g_1^T(z)g_1(z) & \sigma_{12}g_1^T(z)g_1(z)\\ \sigma_{12}g_1^T(z)g_1(z) & \sigma_{22}g_1^T(z)g_1(z)+\frac{1}{\sigma^{22}}g_2^T(z)g_2(z)\end{pmatrix}.$$

This gives

$$h(z,\xi_U)=\prod_{i=1}^{2}V_{ii}(z,\xi_U)=\sigma_{11}\left(1+\sum_{i=1}^{m}z_i^2\right)\left[\sigma_{22}\left(1+\sum_{i=1}^{m}z_i^2\right)+\frac{1}{\sigma^{22}}\sum_{1\le i<j\le m}z_i^2z_j^2\right].$$

Thus, for $L=1$ we have

$$\psi_1(\xi_U)=\int_{\mathcal{Z}}h\left(z,\xi^*\right)\,d\mu(z)=\sigma_{11}\sigma_{22}\frac{45+34m+5m^2}{45}+\frac{\sigma_{11}}{\sigma^{22}}(m^2-m)\left(\frac{11}{90}+\frac{m-2}{54}\right),$$

and

$$\left(\prod_{i=1}^{2}V_{ii}(z,\xi_U)\right)\sum_{i=1}^{2}\frac{F^T(z)e_ie_i^TF(z)}{V_{ii}(z,\xi_U)}=\begin{pmatrix}V_{22}g_1(z)g_1^T(z) & 0 & 0\\ 0 & V_{11}g_1(z)g_1^T(z) & V_{11}g_1(z)g_2^T(z)\\ 0 & V_{11}g_2(z)g_1^T(z) & V_{11}g_2(z)g_2^T(z)\end{pmatrix}.$$

Therefore,

$$Q_1(\xi_U)=\int_{\mathcal{Z}}\left(\prod_{i=1}^{2}V_{ii}(z,\xi_U)\right)\sum_{i=1}^{2}\frac{F^T(z)e_ie_i^TF(z)}{V_{ii}(z,\xi_U)}d\mu(z)=\begin{pmatrix}B_1 & 0 & 0\\ 0 & B_2 & 0\\ 0 & 0 & B_3\end{pmatrix},$$

where

$$B_1=\begin{pmatrix}\frac{(m+3)\sigma_{22}}{3}+\frac{m^2-m}{18\sigma^{22}} & 0\\ 0 & \left(\frac{(5m+9)\sigma_{22}}{45}+\frac{5m^2+3m-8)}{270\sigma^{22}}\right)I_m\end{pmatrix},$$

and

$$B_2=\begin{pmatrix}\frac{(m+3)\sigma_{11}}{3} & 0\\ 0 & \frac{(5m+9)\sigma_{22}}{45}I_m\end{pmatrix},\quad B_3=\frac{(5m+23)\sigma_{11}}{135}I_{(m^2-m)/2}.$$

It follows that

$$\begin{aligned}&F(x)M^{-1}(\xi_U)Q_1(\xi_U)M^{-1}(\xi_U)F^T(x)\\&=\begin{pmatrix}g_1^T(x)(\sigma_{11}^2B_1+\sigma_{12}^2B_2)g_1(x) & g_1^T(x)(\sigma_{11}\sigma_{12}B_1+\sigma_{12}\sigma_{22}B_2)g_1(x)\\ g_1^T(x)(\sigma_{11}\sigma_{12}B_1+\sigma_{12}\sigma_{22}B_2)g_1(x) & g_1^T(x)\left(\sigma_{12}^2B_1+\sigma_{22}^2B_2\right)g_1(x)+\left(\frac{1}{\sigma^{22}}\right)^2g_2^T(x)B_3g_2(x)\end{pmatrix},\end{aligned}$$

and thus

$$\begin{aligned}&\mathrm{tr}\left\{M^{-1}(\xi_U)F^T(x)\Sigma^{-1}F(x)M^{-1}(\xi_U)Q_1(\xi_U)\right\}\\&=\sigma^{11}g_1^T(x)\left(\sigma_{11}^2B_1+\sigma_{12}^2B_2\right)g_1(x)+2\sigma^{12}g_1^T(x)(\sigma_{11}\sigma_{12}B_1+\sigma_{12}\sigma_{22}B_2)g_1(x)\\&\quad+\sigma^{22}g_1^T(x)\left(\sigma_{12}^2B_1+\sigma_{22}^2B_2\right)g_1(x)+\frac{1}{\sigma^{22}}g_2^T(x)B_3g_2(x)\\&=g_1^T(x)(\sigma_{11}B_1+\sigma_{22}B_2)g_1(x)+\frac{1}{\sigma^{22}}g_2^T(x)B_3g_2(x)\\&=\left(\frac{2(5m+9)\sigma_{11}\sigma_{22}}{45}+\frac{(5m^2+3m-8)\sigma_{11}}{270\sigma^{22}}\right)\sum_{i=1}^{m}x_i^2\\&\quad+\frac{(5m+23)\sigma_{11}}{135\sigma^{22}}\sum_{1\le i<j\le m}x_i^2x_j^2+\frac{2(m+3)\sigma_{11}\sigma_{22}}{3}+\frac{(m^2-m)\sigma_{11}}{18\sigma^{22}}\\&\le\left(\frac{2(5m+9)\sigma_{11}\sigma_{22}}{45}+\frac{(5m^2+3m-8)\sigma_{11}}{270\sigma^{22}}\right)m\\&\quad+\frac{(5m+23)\sigma_{11}}{135\sigma^{22}}\frac{m^2-m}{2}+\frac{2(m+3)\sigma_{11}\sigma_{22}}{3}+\frac{(m^2-m)\sigma_{11}}{18\sigma^{22}}\end{aligned}$$

$$
\begin{aligned}
&= 2\left(\frac{\sigma_{11}\sigma_{22}(45+34m+5m^2)}{45} + \frac{(m^2-m)(5m+23)\sigma_{11}}{270\sigma^{22}}\right) \\
&= 2\psi_1(\xi_U).
\end{aligned}
$$

From the above, it is clear that $\phi_1(x,\xi_U)$ is nonnegative for any $x \in [-1,1]^m$, and attains its maximum $r=2$ at $x \in \{(\pm 1, \pm 1, \ldots, \pm 1)\}$, the support points of ξ_U. It follows from Theorem 1′ that the design ξ_U is R_1-optimal over the class Ξ.

For the case $L=\infty$, the answering set corresponding to ξ_U is

$$
\mathcal{A}(\xi_U) = \left\{ z \in \mathcal{Z} \,\middle|\, \prod_{i=1}^{r} V_{ii}(z,\xi_U) = \sup_{z' \in \mathcal{Z}} \prod_{i=1}^{r} V_{ii}(z',\xi_U) \right\} = \{(\pm 1, \pm 1, \ldots, \pm 1)\}.
$$

Take the probability measure $\mu^* = \xi_U$, we have that both $V_{11}(z,\xi_U) = e_1^T V(z,\xi_U) e_1$ and $V_{22}(z,\xi_U) = e_2^T V(z,\xi_U) e_2$ are constants on $\mathcal{A}(\xi_U)$, i.e.,

$$
V_{11}(z,\xi_U) = \sigma_{11}(m+1), \quad V_{22}(z,\xi_U) = \sigma_{22}(m+1) + \frac{m^2-m}{2\sigma^{22}}.
$$

Straightforward calculation gives

$$
\int_{\mathcal{A}(\xi_U)} \frac{F^T(z)e_i e_i^T F(z)}{V_{ii}(z,\xi_U)} d\mu^*(z) = \begin{pmatrix} C_1 & 0 & 0 \\ 0 & C_2 & 0 \\ 0 & 0 & C_3 \end{pmatrix},
$$

where

$$
C_1 = V_{11}^{-1} I_{(m+1)}, \quad C_2 = V_{22}^{-1} I_{(m+1)}, \quad C_3 = V_{22}^{-1} I_{(m^2-m)/2}.
$$

It follows that

$$
\begin{aligned}
\phi_\infty(x,\xi_U) &= \mathrm{tr}\left\{ \Sigma^{-1} F(x) M^{-1}(\xi_U) \left(\int_{\mathcal{Z}} \sum_{i=1}^{2} \frac{F^T(z)e_i e_i^T F(z)}{V_{ii}(z,\xi_U)} d\mu^*(z) \right) M^{-1}(\xi_U) F^T(x) \right\} \\
&= g_1^T(x)(\sigma_{11}C_1 + \sigma_{22}C_2) g_1(x) + \frac{1}{\sigma^{22}} g_2^T(x) C_3 g_2(x) \\
&= \left(\sigma_{11}V_{11}^{-1} + \sigma_{22}V_{22}^{-1}\right)\left(\sum_{i=1}^{m} x_i^2 + 1\right) + \frac{1}{\sigma^{22}} V_{22}^{-1} \sum_{1 \le i < j \le m} x_i^2 x_j^2 \\
&\le \left(\sigma_{11}V_{11}^{-1} + \sigma_{22}V_{22}^{-1}\right)(m+1) + V_{22}^{-1}\frac{m^2-m}{2\sigma^{22}} = 2,
\end{aligned}
$$

where the equality occurs at any support points $x = (\pm 1, \pm 1, \ldots, \pm 1)$. Therefore, the design ξ_U is R_∞-optimal over the class Ξ by Theorem 2.

Example 3. *Parallel linear model with two responses.* In this example we determine the R_1- and R_∞-optimal designs for the following parallel linear model:

$$
\begin{cases} y_1(x) = \theta_{01} + \theta_1 x_1 + \varepsilon_1, \\ y_2(x) = \theta_{02} + \theta_1 x_2 + \varepsilon_2, \end{cases} \tag{15}
$$

where $x = (x_1, x_2) \in \mathcal{X} = \mathcal{Z} = [-1,1]^2$. The variance–covariance matrix Σ of the two responses is assumed to be $\Sigma = (1-\rho)I_2 + \rho J_2$, where I_2 is the 2×2 identity matrix and J_2 is the 2×2 matrix of one's.

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

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

The D-optimal design for estimating θ in this model obtained by Huang et al. (2006) is

$$
\xi_D^* = \begin{Bmatrix} (-1,1) & (1,-1) \\ 1/2 & 1/2 \end{Bmatrix} \quad \text{if } \rho > 0, \tag{16}
$$

and

$$
\xi_D^* = \begin{Bmatrix} (-1,-1) & (1,1) \\ 1/2 & 1/2 \end{Bmatrix} \quad \text{if } \rho < 0. \tag{17}
$$

Now by Theorem 3, we can verify that this D-optimal design is also R_1- and R_∞-optimal for model (15). Only the case $\rho > 0$ is shown below. The case $\rho < 0$ can be treated in the similar way.

First, find the set $\mathcal{R}_3$ defined in (9). 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 $\|K\|=1$. Letting

$$a=\frac{1}{2\sqrt{1+\rho}}+\frac{1}{2\sqrt{1-\rho}},\quad b=\frac{1}{2\sqrt{1+\rho}}-\frac{1}{2\sqrt{1-\rho}},$$

and a straightforward calculation gives

$$H(x)=\begin{pmatrix} aK_{11}+bK_{21} & aK_{12}+bK_{22} & aK_{13}+bK_{23} \\ bK_{11}+aK_{21} & bK_{12}+aK_{22} & bK_{13}+aK_{23} \\ H_{31}(x) & H_{32}(x) & H_{33}(x) \end{pmatrix},$$

where

$$\begin{aligned} H_{31}(x)&=(ax_1+bx_2)K_{11}+(bx_1+ax_2)K_{21}=x_1H_{11}+x_2H_{21},\\ H_{32}(x)&=(ax_1+bx_2)K_{12}+(bx_1+ax_2)K_{22}=x_1H_{12}+x_2H_{22},\\ H_{33}(x)&=(ax_1+bx_2)K_{13}+(bx_1+ax_2)K_{23}=x_1H_{13}+x_2H_{23}. \end{aligned}$$

Note from

$$\begin{pmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \end{pmatrix}=\Sigma^{-1/2}K$$

and $\|K\|^2=\mathrm{tr}(K^TK)=1$, we have

$$\sum_{i=1}^{2}\sum_{j=1}^{3}H_{ij}^2+2\rho\sum_{j=1}^{3}H_{1j}H_{2j}=1.$$

For $\rho>0$, it is not difficult to obtain

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

and the boundary of $\mathcal{R}_3$ is obtained from the points $x=(-1,1)$ and $x=(1,-1)$. Therefore, the support points $(-1,1)$ and $(1,-1)$ of ξ_D^* are also the support points of the R_1-optimal design in the case $\rho>0$. We take

$$\gamma=\left\{\int_{\mathcal{Z}}\prod_{i=1}^{2}V_{ii}(z,\xi_D^*)\,d\mu(z)\right\}^{-1/2}=\frac{6}{7-\rho},$$

$$A(\xi_D^*)=\sqrt{\frac{7-\rho}{3}}\mathrm{diag}\left(\frac{1}{2},\frac{1}{2},\frac{\sqrt{6}}{6}\right),$$

and

$$D=\gamma M^{-1}(\xi_D^*)A(\xi_D^*)=\frac{1}{2\sqrt{7-\rho}}\begin{pmatrix} 2\sqrt{3} & 2\rho\sqrt{3} & 0 \\ 2\rho\sqrt{3} & 2\sqrt{3} & 0 \\ 0 & 0 & \sqrt{2}(1-\rho) \end{pmatrix}.$$

Corresponding to the two support points $(1,-1)$ and $(-1,1)$, we take

$$\begin{aligned} K_1&=\Sigma^{-1/2}F(1,-1)D\\ &=\frac{1}{2\sqrt{7-\rho}}\begin{pmatrix} \sqrt{3}(\sqrt{1+\rho}+\sqrt{1-\rho}) & \sqrt{3}(\sqrt{1+\rho}-\sqrt{1-\rho}) & \sqrt{2}\sqrt{1-\rho} \\ \sqrt{3}(\sqrt{1+\rho}-\sqrt{1-\rho}) & \sqrt{3}(\sqrt{\rho+1}+\sqrt{1-\rho}) & -\sqrt{2}\sqrt{1-\rho} \end{pmatrix}, \end{aligned}$$

and

$$\begin{aligned} K_2&=\Sigma^{-1/2}F(-1,1)D\\ &=\frac{1}{2\sqrt{7-\rho}}\begin{pmatrix} \sqrt{3}(\sqrt{1+\rho}+\sqrt{1-\rho}) & \sqrt{3}(\sqrt{1+\rho}-\sqrt{1-\rho}) & -\sqrt{2}\sqrt{1-\rho} \\ \sqrt{3}(\sqrt{1+\rho}-\sqrt{1-\rho}) & \sqrt{3}(\sqrt{\rho+1}+\sqrt{1-\rho}) & \sqrt{2}\sqrt{1-\rho} \end{pmatrix}. \end{aligned}$$

It can be verified that conditions (i), (ii) and (iii) of Theorem 3 are satisfied. Therefore, the design ξ_D^* for $\rho>0$ is also R_1-optimal for model (15).

Similarly, for $L=\infty$, it is easy to verify that ξ_D^* for $\rho>0$ satisfies the conditions in Theorem 3 by taking

$$\mu^*=\left\{\begin{array}{cccc} (1,-1) & (-1,1) & (1,1) & (-1,-1) \\ 1/4 & 1/4 & 1/4 & 1/4 \end{array}\right\},$$

and $\gamma=1$,

$$A(\xi^*_\infty)=\frac{1}{\sqrt{3-\rho}}\begin{pmatrix}1&0&0\\0&1&0\\0&0&\sqrt{2}\end{pmatrix},\quad D=\frac{1}{\sqrt{3-\rho}}\begin{pmatrix}1&\rho&0\\\rho&1&0\\0&0&\frac{\sqrt{2}(1-\rho)}{2}\end{pmatrix},$$

$$K_1=\frac{1}{2\sqrt{3-\rho}}\begin{pmatrix}\sqrt{1+\rho}+\sqrt{1-\rho}&\sqrt{1+\rho}-\sqrt{1-\rho}&\sqrt{2}\sqrt{1-\rho}\\\sqrt{1+\rho}-\sqrt{1-\rho}&\sqrt{1+\rho}+\sqrt{1-\rho}&-\sqrt{2}\sqrt{1-\rho}\end{pmatrix},$$

$$K_2=\frac{1}{2\sqrt{3-\rho}}\begin{pmatrix}\sqrt{1+\rho}+\sqrt{1-\rho}&\sqrt{1+\rho}-\sqrt{1-\rho}&-\sqrt{2}\sqrt{1-\rho}\\\sqrt{1+\rho}-\sqrt{1-\rho}&\sqrt{1+\rho}+\sqrt{1-\rho}&\sqrt{2}\sqrt{1-\rho}\end{pmatrix}.$$

Therefore, the design ξ^*_D for $\rho>0$ is also R_∞-optimal for model (15).

5. Concluding remarks

The new optimality for designing multiresponse experiments are studied in this paper. The R^r_L-optimal design minimizes the volume of the r-dimensional rectangular confidence region for predicting the r responses in the multiresponse linear regression model. We established a general equivalence theorem and a geometrical characterization of the R^r_L-optimal design, which are useful for obtaining or verifying the R^r_L-optimality of designs. The analytic solutions in three examples are provided which are found by the general equivalence theorem and the geometrical characterization of the R^r_L-optimal designs. It is to be noted that the R^r_L-optimal design must be calculated numerically in most cases. In general, the R^r_L-optimal design may highly depend on the covariance matrix of the responses.

Note also that the D-optimal design minimizes the volume of the ellipsoid of concentration for the vector of unknown parameters, and the I^r_L-optimal design minimizes the volume of the confidence ellipsoid of the r-dimensional response vector. The loss of efficiency might be moderate or substantial, if a rectangular prediction confidence region is constructed on the basis of a D-optimal design or an I^r_L-optimal design, or a prediction confidence ellipsoid is constructed on the basis of a D-optimal design or an R^r_L-optimal design. However, in some particular multiresponse models, some of the D-, R^r_L- and I^r_L-optimal designs coincide.

Appendix A

A.1. Proof of Lemma 2

(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

$$\frac{d}{d\alpha}\prod_{i=1}^{r}V_{ii}(z,\xi_\alpha)=\sum_{i=1}^{r}\left(\prod_{\substack{j=1\\j\neq i}}^{r}V_{jj}(z,\xi_\alpha)\right)e_i^TF(z)M^{-1}(\xi_\alpha)\big(M(\xi)-M(\bar{\xi})\big)M^{-1}(\xi_\alpha)F^T(z)e_i.$$

For all $L\in[1,\infty)$,

$$\begin{aligned}\Delta_{\psi_L}(\xi,\bar{\xi})&=\lim_{\alpha\to0^+}\frac{d\psi_L(\xi_\alpha)}{d\alpha}\\&=\psi_L^{1-L}(\xi)\int_{\mathcal{Z}}\left(\prod_{i=1}^{r}V_{ii}^{L-1}\right)\left\{\sum_{i=1}^{r}\left(\prod_{\substack{j=1\\j\neq i}}^{r}V_{jj}\right)e_i^TFM^{-1}(\xi)\big(M(\xi)-M(\bar{\xi})\big)M^{-1}(\xi)F^Te_i\right\}d\mu\\&=r\psi_L(\xi)-\psi_L^{1-L}(\xi)\int_{\mathcal{Z}}\left(\prod_{i=1}^{r}V_{ii}^{L}\right)\sum_{i=1}^{r}\frac{e_i^TFM^{-1}(\xi)M(\bar{\xi})M^{-1}(\xi)F^Te_i}{V_{ii}}d\mu\\&=r\psi_L(\xi)-\psi_L^{1-L}(\xi)\ \mathrm{tr}\{M^{-1}(\xi)M(\bar{\xi})M^{-1}(\xi)Q_L(\xi)\}.\end{aligned}$$

(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.

A.2. Proof of Theorem 3

We present the proof for the case of $L\in[1,\infty)$, while the case of $L=\infty$ can be treated similarly. For $L\in[1,\infty)$ let $\xi=[x_v,w_v]_{v=1}^s\in\Xi$ denote an R_L^r-optimal design. By Theorem 1′ we have

$$\phi_L(x,\xi)=\frac{\mathrm{tr}\{M^{-1}(\xi)F^T(x)\Sigma^{-1}F(x)M^{-1}(\xi)Q_L(\xi)\}}{\int_{\mathcal{Z}}\prod_{i=1}^r V_{ii}^L(z,\xi)\,d\mu(z)}\le r \tag{18}$$

for all $x\in\mathcal{X}$, and

$$\phi_L(x_v,\xi)=\frac{\mathrm{tr}\{M^{-1}(\xi)F^T(x_v)\Sigma^{-1}F(x_v)M^{-1}(\xi)Q_L(\xi)\}}{\int_{\mathcal{Z}}\prod_{i=1}^r V_{ii}^L(z,\xi)\,d\mu(z)}=r \tag{19}$$

for all $v=1,\dots,s$. Let

$$\gamma=\left(\int_{\mathcal{Z}}\prod_{i=1}^r V_{ii}^L(z,\xi)\,d\mu(z)\right)^{-1/2}\quad\text{and}\quad D=\gamma M^{-1}(\xi)A(\xi).$$

It follows that

$$\gamma A(\xi)=M(\xi)D=\sum_{v=1}^s w_vF^T(x_v)\Sigma^{-1/2}K_v,$$

where $K_v=\Sigma^{-1/2}F(x_v)D$, $v=1,\dots,s$. This proves the representation given in (i).

Eq. (19) and the representation of $\gamma A(\xi)$ yield

$$\begin{aligned}\|K_v\|^2&=\mathrm{tr}\left\{K_v^TK_v\right\}\\&=\mathrm{tr}\left\{D^TF^T(x_v)\Sigma^{-1}F(x_v)D\right\}\\&=\mathrm{tr}\left\{\gamma^2A(\xi)^TM^{-1}(\xi)F^T(x_v)\Sigma^{-1}F(x_v)M^{-1}(\xi)A(\xi)\right\}\\&=\mathrm{tr}\left\{\gamma^2M^{-1}(\xi)F^T(x_v)\Sigma^{-1}F(x_v)M^{-1}(\xi)A(\xi)A(\xi)^T\right\}\\&=\frac{\phi_L(x_v,\xi)}{r}=1,\end{aligned}$$

and shows condition (iii).

From the inequality (18) and the Cauchy–Schwarz inequality we get

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

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

$$\begin{aligned}\mathrm{tr}\left\{D^T\gamma A(\xi)\right\}&=\mathrm{tr}\left\{D^T\sum_{v=1}^s w_vF^T(x_v)\Sigma^{-1}F(x_v)D\right\}\\&=\sum_{v=1}^s w_v\,\mathrm{tr}\left\{D^TF^T(x_v)\Sigma^{-1}F(x_v)D\right\}\\&=\sum_{v=1}^s w_v\,\mathrm{tr}\left\{\gamma^2A(\xi)^TM^{-1}(\xi)F^T(x_v)\Sigma^{-1}F(x_v)M^{-1}(\xi)A(\xi)\right\}\\&=\sum_{v=1}^s w_v\,\mathrm{tr}\left\{\gamma^2M^{-1}(\xi)F^T(x_v)\Sigma^{-1}F(x_v)M^{-1}(\xi)A(\xi)A(\xi)^T\right\}\\&=\sum_{v=1}^s w_v\frac{\phi_L(x_v,\xi)}{r}\\&=\sum_{v=1}^s w_v=1.\end{aligned}$$

Therefore, $\gamma A(\xi)$ is a boundary point of $\mathcal{R}_p$ with supporting hyperplane D which proves (ii).

To prove sufficiency let D denote a supporting hyperplane to the convex hull $\mathcal{R}_p$ at the boundary point $\gamma A(\xi)$. Thus we have for all $x\in\mathcal{X}$ and K satisfying $\|K\|=1$

$$|\mathrm{tr}\{D^TF^T(x)\Sigma^{-1/2}K\}|\le 1. \tag{20}$$

Especially, by taking $K=\Sigma^{-1/2}F(x)D/\sqrt{\mathrm{tr}\{D^TF^T(x)\Sigma^{-1}F(x)D\}}$, (20) implies

$$\mathrm{tr}\{D^TF^T(x)\Sigma^{-1}F(x)D\}\le 1\quad\text{for all }x\in\mathcal{X}. \tag{21}$$

Because D is a supporting hyperplane to $\mathcal{R}_p$ at the boundary point $\gamma A(\xi)$ we obtain from (20) (used at $x=x_v$) and the representation (i) that

$$1=\mathrm{tr}\{D^T\gamma A(\xi)\}=\mathrm{tr}\left\{D^T\sum_{v=1}^{s}w_vF^T(x_v)\Sigma^{-1/2}K_v\right\}=\sum_{v=1}^{s}w_v\,\mathrm{tr}\{D^TF^T(x_v)\Sigma^{-1/2}K_v\}\le 1$$

and this implies $\mathrm{tr}\{D^TF^T(x_v)\Sigma^{-1/2}K_v\}=1$, $v=1,\ldots,s$. By an application of the Cauchy–Schwarz inequality we now get for $v=1,\ldots,s$

$$1=(\mathrm{tr}\{D^TF^T(x_v)\Sigma^{-1/2}K_v\})^2\le\mathrm{tr}\{K_v^TK_v\}\mathrm{tr}\{D^TF^T(x_v)\Sigma^{-1}F(x_v)D\}\le 1, \tag{22}$$

where the last inequality results from (21) and condition (iii). Therefore we have $K_v=\lambda_v\Sigma^{-1/2}F(x_v)D$ for some $\lambda_v\in\mathbb{R}$, $v=1,\ldots,s$. From the normalizing conditions on the K_v in (iii) we thus obtain

$$1=\mathrm{tr}\{K_v^TK_v\}=\lambda_v^2\,\mathrm{tr}\{D^TF^T(x_v)\Sigma^{-1}F(x_v)D\}=\lambda_v^2,\quad v=1,\ldots,s. \tag{23}$$

On the other hand, we have from the property that $\gamma A(\xi)$ is a boundary point of $\mathcal{R}_p$ with supporting hyperplane D

$$\begin{aligned}1=\mathrm{tr}\{D^T\gamma A(\xi)\}&=\sum_{v=1}^{s}w_v\,\mathrm{tr}\{D^TF^T(x_v)\Sigma^{-1/2}K_v\}\\&=\sum_{v=1}^{s}w_v\lambda_v\,\mathrm{tr}\{D^TF^T(x_v)\Sigma^{-1}K_v\}=\sum_{v=1}^{s}w_v\lambda_v.\end{aligned}$$

Eq. (23) and $w_v>0$ with $\sum_{v=1}^{s}w_v=1$ now show that $\lambda_v=1$, which implies $K_v=\Sigma^{-1/2}F(x_v)D$, $v=1,\ldots,s$. From this representation we finally obtain

$$\gamma A(\xi)=\sum_{v=1}^{s}w_vF^T(x_v)\Sigma^{-1/2}K_v=\sum_{v=1}^{s}w_vF^T(x_v)\Sigma^{-1}F(x_v)D=M(\xi)D.$$

It follows that

$$\begin{aligned}1&=\mathrm{tr}\left\{D^T\gamma A(\xi)\right\}\\&=\gamma^2\,\mathrm{tr}\left\{M^{-1}(\xi)A(\xi)A(\xi)^T\right\}\\&=\frac{\gamma^2}{r}\mathrm{tr}\left\{M^{-1}(\xi)Q_L(\xi)\right\}\\&=\frac{\gamma^2}{r}\int_{\mathcal{Z}}\left(\prod_{i=1}^{r}V_{ii}^L(z,\xi)\right)\sum_{i=1}^{r}\frac{e_i^TF(z)M^{-1}(\xi)F^T(z)e_i}{V_{ii}(z,\xi)}d\mu(z)\\&=\frac{\gamma^2}{r}\int_{\mathcal{Z}}\left(\prod_{i=1}^{r}V_{ii}^L(z,\xi)\right)\sum_{i=1}^{r}\frac{e_i^TV(z,\xi)e_i}{V_{ii}(z,\xi)}d\mu(z)\\&=\gamma^2\int_{\mathcal{Z}}\left(\prod_{i=1}^{r}V_{ii}^L(z,\xi)\right)d\mu(z),\end{aligned}$$

and the inequality (21) yields

$$\begin{aligned}\phi_L(x,\xi)&=\frac{\mathrm{tr}\{M^{-1}(\xi)F^T(x)\Sigma^{-1}F(x)M^{-1}(\xi)Q_L(\xi)\}}{\int_{\mathcal{Z}}\prod_{i=1}^{r}V_{ii}^L(z,\xi)\,d\mu(z)}\\&=\frac{r\,\mathrm{tr}\{M^{-1}(\xi)F^T(x)\Sigma^{-1}F(x)M^{-1}(\xi)A(\xi)A(\xi)^T\}}{\int_{\mathcal{Z}}\prod_{i=1}^{r}V_{ii}^L(z,\xi)\,d\mu(z)}\\&=\frac{r\,\mathrm{tr}\{A(\xi)^TM^{-1}(\xi)F^T(x)\Sigma^{-1}F(x)M^{-1}(\xi)A(\xi)\}}{\int_{\mathcal{Z}}\prod_{i=1}^{r}V_{ii}^L(z,\xi)\,d\mu(z)}\\&=r\,\mathrm{tr}\{D^TF^T(x)\Sigma^{-1}F(x)D\}\le r\end{aligned}$$

for all $x\in\mathcal{X}$. By an application of Theorem 1′ it now follows that the design ξ is R_L^r-optimal, which completes the proof of Theorem 3.

References

Atashgah, A.B., Seifi, A., 2009. Optimal design of multi-response experiments using semi-definite programming. Optimization and Engineering 10, 75–90.
Chang, S.I., 1994. Some properties of multi-response D-optimal designs. Journal of Mathematical Analysis and Applications 184, 256–262.
Danskin, J.M., 1967. The Theory of Max–Min and its Application to Weapons Allocation Problems. Springer-Verlag, New York.
Dette, H., 1993a. Elfving's theorem for D-optimality. Annals of Statistics 21, 753–766.
Dette, H., 1993b. A new interpretation of optimality for E-optimal designs in linear regression models. Metrika 40, 37–50.
Dette, H., 1993c. Bayesian D-optimal and model robust designs in linear regression models. Statistics 25, 27–46.
Dette, H., 1996. A note on Bayesian C- and D-optimal designs in nonlinear regression models. Annals of Statistics 24, 1225–1234.

Dette, H., 1997. Designing experiments with respect to ‘standardized’ optimality criteria. Journal of the Royal Statistical Society. Series B 59, 97–110.

Dette, H., O'Brien, T.E., 1999. Optimality criteria for regression models based on predicted variance. Biometrika 86, 93–106.

Draper, N.R., 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. Annals of Mathematical Statistics 23, 255–262.

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

Gaffke, H., Heiligers, B., 1996. Approximate designs for polynomial regression: invariance, admissibility, and optimality. In: Ghosh, S., Rao, C.R. (Eds.), Handbook of Statistics: Design and Analysis of Experiments. Elsevier, Amsterdam, pp. 1149–1199.

Gueorguieva, I., Aarons, L., Ogungbenro, K., Jorga, K.M., Rodgers, T., Rowland, M., 2006. Optimal design for multivariate response pharmacokinetic models. Journal of Pharmacokinetics and Pharmacodynamics 33, 97–123.

Holland-Letz, T., Dette, H., Pepelyshev, A., 2009. A geometric characterization of optimal designs for regression models with correlated observations. Journal of the Royal Statistical Society. Series B 73, 239–252.

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

Jin, H.-D., Yue, R.-X., 2008. D- and A-optimal designs for mixture experiments with multiresponse models. Journal of Shanghai Normal University 37, 124–130.

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

Krafft, O., Schaefer, M., 1992. D-optimal designs for a multivariate regression model. Journal of Multivariate Analysis 42, 130–140.

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

Sagnol, G., 2011. Computing optimal designs of multiresponse experiments reduces to second-order cone programming. Journal of Statistical Planning and Inference 141, 1684–1708.

Studden, W.J., 1971. Elfving's theorem and optimal designs for quadratic loss. Annals of Mathematical Statistics 42, 1613–1621.

Wang, Q.-J., Yue, R.-X., 2008. Equivalent theorem of Bayesian optimal design for multiresponse linear regression models. Journal of Shanghai Normal University 37, 229–233.

Whittle, P., 1973. Some general points in the theory of optimal experimental designs. Journal of the Royal Statistical Society. Series B 35, 123–130.

Yue, R.-X., Liu, X., 2010. I_L^r-optimal designs for a hierarchically ordered system of regression models. Computational Statistics & Data Analysis 54, 3458–3465.

Zhu, J., Yue, R.-X., Wu, J.-W., 2010. Robust designs against outliers for multiresponse models. Journal of Shanghai Normal University 39, 1–6.