OPTIMALITY REGIONS FOR DESIGNS IN MULTIPLE LINEAR REGRESSION MODELS WITH CORRELATED RANDOM COEFFICIENTS

ULRIKE GRASSHOFF, HEINZ HOLLING, FRANK RÖTTGER, AND RAINER SCHWABE

Abstract. This paper studies optimal designs for linear regression models with correlated effects for single responses. We introduce the concept of rhombic design to reduce the computational complexity and find a semi-algebraic description for the $D$-optimality of a rhombic design via the Kiefer–Wolfowitz equivalence theorem. Subsequently, we show that the structure of an optimal rhombic design depends directly on the correlation structure of the random coefficients.

Keywords. $D$-optimal design, heteroscedastic model, random coefficients, multiple regression, semi-algebraic geometry

1. Introduction

Hierarchical regression models with random coefficients enjoy growing importance in biological and psychological applications, whenever there is a variation with respect to the observed subjects. Hereby we often cannot expect the random coefficients to be uncorrelated, which means that we assume that the random coefficients are e.g. normally distributed with a population mean and a non-diagonal covariance matrix. A special model that will be the topic of this paper are random effects models for linear regression with singular responses, which means that we obtain only one observation per unit. This particular model was motivated by Freund and Holling in [4] and Patan and Bogacka in [8]. A natural question that arises is to find optimal experimental designs for these models with respect to some optimality criterion. Graßhoff et al. determined $D$-optimal designs that maximize the determinant of the corresponding information matrix, for a couple of different covariance structures in [6] and [5]. They found that in contrast to fixed effects models for multiple linear regression optimal settings may, surprisingly, occur in the interior of the design region under certain conditions on the covariance structure of the random coefficients. In the present paper, we investigate conditions on the covariance structure to discriminate situations in which optimal designs are completely supported on the boundary of the design region as in fixed effects models and situations in which optimal designs may have additional support points in the interior. This is done for the special class of rhombic designs, which are invariant with respect to permutations of the regressors and simultaneous sign change and which we will introduce in Section 3. Section 4 shows via the Kiefer-Wolfowitz equivalence theorem [9, Theorem 3.7] how the parameter regions for which rhombic designs with or without interior points are $D$-optimal are described by semi-algebraic sets, which are sets defined by polynomial inequalities and equations and how the optimality depends on the covariance structure. Furthermore, we show that for the assumed covariance structure of the random coefficients, the $D$-optimality of designs with interior support points translates to a simple matrix equation for the information. We show as a consequence of the results in Section 4 that the distinction, whether a $D$-optimal rhombic design requires interior support points or not, can be made by evaluating a polynomial only dependent on the covariance matrix of the random coefficients. Based on these results, we are able to compute

Corresponding author: Frank Röttger.
optimal designs and their optimality regions explicitly for small to moderate dimensions in Section 5 and we conjecture results for arbitrary dimensions in Section 6.

2. General setup

We consider a random coefficient regression model \( Y_i(x_i) = f(x_i)^T b_i + \varepsilon_i, \) \( i = 1, \ldots, n \) for observations \( Y_i \) at experimental settings \( x_i \) where \( f \) is a \( p \)-dimensional vector of linearly independent regression functions, \( b_i \) is a \( p \)-dimensional vector of random coefficients and \( \varepsilon_i \) are additional observational errors. The random coefficients are assumed to be distributed with unknown mean vector \( \beta \) and prespecified dispersion matrix \( D \), whereas the error terms \( \varepsilon_i \) are distributed with zero mean and equal variance \( \sigma^2 \). Moreover the random coefficients and the error terms are assumed to be uncorrelated. In this note we assume that all observations \( Y_i \) are independent, i.e. only one observation is made for each realization \( b_i \) of the random coefficients. Moreover, we assume here that an intercept is included in the model \( (f_1(x) \equiv 1) \) such that the additive observational error \( \varepsilon_i \) may be subsumed into the random intercept. This can be achieved by substituting the first entry \( b_{i1} \) in the random coefficient vector by \( b_{i1} + \varepsilon_i \) and the first entry \( d_0 \) in the dispersion matrix \( D \) by \( d_0 + \sigma^2 \). The model can hence be rewritten as a heteroscedastic linear fixed effects model,

\[
Y_i(x_i) = f(x_i)^T \beta + \varepsilon_i,
\]

where now \( \varepsilon_i = f(x_i)^T (b_i - \beta) \) with mean zero and the variance function defined by \( \sigma^2(x) = f(x)^T D f(x) \). Within this heteroscedastic linear model for each single setting \( x \) in a design region \( \mathcal{X} \) the elemental information matrix \( [1] \) equals \( M(x) = f(x)^T f(x)/\sigma^2(x) \), assuming that \( \sigma^2(x) > 0 \) for all \( x \in \mathcal{X} \). An exact design of sample size \( n \) can be described by a finite set of mutually distinct settings \( x_j, j = 1, \ldots, m, \) for the explanatory variable and the corresponding numbers \( n_j \) of replications at \( x_j \), where \( x_j \) may be chosen from the design region \( \mathcal{X} \) of potential settings. Equivalently, a standardized version \( \xi \) may be characterized by the proportions \( \xi(x_j) = n_j/n \) at settings \( x_j \). We call \( \xi(x_j) \) the design weight at design point \( x_j \). Then for a design \( \xi \), the standardized (per observation) information matrix is given by \( M(\xi) = \sum_{j=1}^m \xi(x_j) M(x_j) \), which is proportional to the finite sample information matrix with a normalizing constant \( \frac{1}{n} \). Note that the covariance matrix for the weighted least squares estimator \( \hat{\beta} \), which is the best linear unbiased estimator for \( \beta \) is proportional to the inverse of the information matrix. Hence, maximizing the information matrix is equivalent to minimizing the covariance matrix of \( \hat{\beta} \).

To compare different designs we consider the most popular criterion, the \( D \)-criterion, with respect to which a design \( \xi \) is \( D \)-optimal, if it maximizes the logarithm \( \log \det(M(\xi)) \) of the determinant of the information matrix. This is equivalent to the minimization of the volume of the confidence ellipsoid for \( \beta \) under the assumption of normally distributed errors. As discrete optimization on the set of exact designs is generally too complicated we relax the condition on the weights \( \xi(x_j) \) being multiples of \( \frac{1}{n} \) and consider approximate designs \( \xi \) in the spirit of Kiefer [7] with real-valued weights \( w_j = \xi(x_j) \geq 0 \) satisfying \( \sum_{j=1}^n w_j = 1 \). A detailed introduction to the theory of optimal design is [9].

In the setting of approximate designs, for which the proportions \( \xi(x) \) are not necessarily multiples of \( 1/n \), where \( n \) denotes the sample size, the \( D \)-optimality of a design \( \xi^* \) can be established by the well-known Kiefer-Wolfowitz equivalence theorem (see [2, Theorem 2.2.1], for a suitable version):

**Theorem 2.1** (Extended Kiefer-Wolfowitz equivalence theorem). A design \( \xi^* \) is \( D \)-optimal on \( \mathcal{X} \), if and only if \( f(x)^T M(\xi^*)^{-1} f(x)/\sigma^2(x) \leq p, \) uniformly in \( x \in \mathcal{X} \).

Let \( \psi(x; \xi) := f(x)^T (pD - M(\xi)^{-1}) f(x) \) be the suitably transformed sensitivity function. When we substitute \( \sigma^2(x) = f(x)^T D f(x) \) into this relation and rearrange terms, \( D \)-optimality is achieved,
if
\begin{equation}
\psi(x; \xi^*) \geq 0
\end{equation}
for all \( x \in \mathcal{X} \). Moreover, equality is attained in (2.2) for design points in the support of an optimal design \( \xi^* \):

**Corollary 2.2** ([9, Corollary 3.10]). It holds that \( \psi(x; \xi^*) = 0 \) for all \( x \in \mathcal{X} \) with \( \xi^*(x) > 0 \).

For notational convenience we define
\[ \Gamma(\xi) := pD - M(\xi)^{-1}, \]
such that \( \psi(x; \xi) = f(x)^T \Gamma(\xi) f(x) \).

### 3. Multiple linear regression

In the following we consider the situation of a multiple linear regression model with \( K \) factors where we have \( n \) observations
\begin{equation}
Y_i(x_i) = \beta_0 + \sum_{k=1}^{K} \beta_k x_{ik} + \varepsilon_i
\end{equation}
with \( x_i = (x_{i1}, ..., x_{iK})^T \in \mathcal{X} = [-1, 1]^K \) and \( \text{Var}(\varepsilon_i) = \sigma^2(x_i) = f(x_i)^T D f(x_i) \). Here we assume that we can choose the design points from the symmetric standard hypercube. The vector of regression functions is given by \( f(x) = (1, x_1, ..., x_K)^T \), such that the model contains an intercept by the first component of \( f \). Note that now \( p = K + 1 \) and \( \beta = (\beta_0, ..., \beta_K)^T \).

We assume that the random coefficients \( b_{i1}, ..., b_{iK} \) associated with the components \( x_1, ..., x_K \) of the regressor are homoscedastic with variance \( d_1 \) and equi-correlated with covariance \( d_2 \). Moreover, let the random intercept \( b_0 \) be uncorrelated with the other random coefficients. To be more precise we consider a \( p \times p \)-dimensional dispersion matrix
\begin{equation}
D = \begin{pmatrix} d_0 & 0 \\ 0 & D_1 \end{pmatrix},
\end{equation}
where \( D_1 \) is a completely symmetric \( K \times K \)-dimensional matrix
\[ D_1 = (d_1 - d_2) \mathbb{I}_K + d_2 \mathbb{1}_K \mathbb{1}_K^T. \]
Here, \( \mathbb{I}_K \) defines the \( K \times K \)-dimensional identity matrix and \( \mathbb{1}_K \) the vector of length \( K \) where all entries equal 1.

**Definition 3.1** (Model cone). We define
\[ C_K := \left\{ (d_0, d_1, d_2)^T \in \mathbb{R}^3 \mid d_0 > 0, d_1 > 0, -\frac{d_1}{K-1} \leq d_2 \leq d_1 \right\} \]
as the *model cone*, so the values of \( (d_0, d_1, d_2)^T \) where \( D \) is non-negative definite.

#### 3.1. Diagonal dispersion matrix

To start we first assume additionally that \( d_2 = 0 \) which means that all components of the random coefficients are uncorrelated, so that the dispersion matrix
\[ D = \begin{pmatrix} d_0 & 0 \\ 0 & d_1 \mathbb{I}_K \end{pmatrix} \]
of the random coefficients \( b_i \) is a diagonal matrix. Hence, the variance of each design point is equal to \( \sigma^2(x) = d_0 + d_1 \sum_{j=1}^{K} x_j^2 \). In [5] it is shown that uniform full factorial \( 2^K \)-designs supported on the points \( (\pm x_1, ..., \pm x_K) \) are \( D \)-optimal. It holds that \( x_1 = \ldots = x_K = x^* \) is optimal and it depends on the values of \( d_0 \) and \( d_1 \) if \( x^* = 1 \) or \( x^* < 1 \). The designs constitute the orbit generated by \( (x_1, ..., x_K) \)
with respect to the (finite) group of transformations of both sign changes within the factors and permutations of the factors themselves. For more details see [5].

3.2. Non-diagonal dispersion matrix. We now assume that \( d_2 \neq 0 \). To reduce the complexity of the system of polynomial equations and inequalities given by the equivalence theorem, one can apply various methods. One of these methods is to assume a certain design structure. A standard approach is the assumption of symmetry in the design under some group action and a restriction of the design region. This is motivated from the symmetry of \( D \)-optimal designs for the situation with \( d_2 = 0 \) as described above in Section 3.1. This applies as the \( D \)-criterion \( \log \det(M(\xi)) \) is not affected by these transformations \( g \) as \( \det M(\xi^g) = \det M(\xi) \). Hence, by convexity the class of invariant designs constitutes an essentially complete class such that search may be restricted to invariant designs. A particular class of invariant designs are rhombic designs. Let \( \text{Sym}(K) \) denote the permutation group on \( K \) elements and \( \{\pm 1\} \) the permutation group with respect to a global sign change, which is therefore isomorphic to \( \text{Sym}(2) \).

**Definition 3.2.** Let a design \( \xi \) for the given model with support on the space diagonals without the origin and at most two points per space diagonal be a **rhombic design** if it is invariant under the action of \( \text{Sym}(K) \times \{\pm 1\} \) on the design points.

This means that we study designs on \( [-1,1]^K \), that are invariant under permutations among the entries of each design point and a global sign change with support on the diagonals of maximal length without the origin. There are \( \lfloor \frac{K}{2} \rfloor + 1 \) different orbits under this group action, where \( \lfloor z \rfloor \) denotes the integer part of some \( z \in \mathbb{R} \). We define \( \tilde{K} := \lfloor \frac{K}{2} \rfloor \). Now, let \( x_\ell \) for \( \ell \in \{0,1,\ldots,\tilde{K}\} \) denote the location parameter of each orbit \( O_\ell(x_\ell) \) with \( 0 < x_\ell \leq 1 \) and either \( \ell \) or \( K - \ell \) negative signs. The location parameter denotes the absolute value of the entries of the design points in \( O_\ell(x_\ell) \) as the design points of a rhombic design are restricted to the space diagonals. Let \( N_\ell = 2\binom{K}{\ell} \) for \( \ell \neq \frac{K}{2} \) and \( N_{\frac{K}{2}} = \binom{K}{\frac{K}{2}} \) for \( \ell = \frac{K}{2} \).

Rhombic designs can be characterized as follows: Let \( x_j \) for \( j \in \{0,1,\ldots,\tilde{K}\} \) be a design point with entries of the same absolute value, i.e. \( x_j \) lies on a space diagonal of \( \mathcal{X} \). Let \( O_j \) be the orbit of \( x_j \) under the action of \( \text{Sym}(K) \times \{\pm 1\} \) and let \( \tilde{x}_j \) be the uniform design on \( O_j \) that assigns the proportion \( \frac{1}{N_j} \) to each \( x \in O_j \). If every orbit \( O_j \) is attributed with a weight \( w_j \geq 0 \) such that \( \sum_{j=0}^{\tilde{K}} w_j = 1 \), then \( \sum_{j=0}^{\tilde{K}} w_j \tilde{x}_j \) is a rhombic design.

To formalize the invariance considerations, let \( g \) denote the group action that generates rhombic designs. Then, there exists a matrix \( Q_g \) so that \( f(g(x)) = Q_g f(x) \).

Figures 1 and 2 exemplify the two different rhombic design classes that will be studied separately. This distinction is made on the location of the design points. With rhombic vertex designs we refer to rhombic designs, where the support is restricted to the vertices of the hypercube, while non-vertex designs are allowed to have points on both the vertices and the interior or in the interior only. In Figure 1, the blue points denote the orbit of \((x_0,x_0)\), with \( 0 < x_0 \leq 1 \), while the red points denote the orbit of \((x_1,-x_1)\), with \( 0 < x_1 \leq 1 \). Similarly, in Figure 2, the blue points denote the orbit of \((x_0,x_0,x_0)\), with \( 0 < x_0 \leq 1 \), while the red points denote the orbit of \((x_1,x_1,-x_1)\), with \( 0 < x_1 \leq 1 \). We chose the name **rhombic designs** due to the structure of the design points in Figure 1.

The usefulness of rhombic design is mainly due to the complexity reduction that comes from its definition: Instead of finding \( K + 1 \) design points each with \( K \) location variables for the entries and a variable for the design weight, we restrict the problem to \( K + 1 \) orbits with one weight variable and one location variable per orbit.

**Lemma 3.3.** The variance function \( \sigma^2(x) = f(x)^T D f(x) \) is equal for all \( x \) in one orbit.
Proof. Let $g$ be the group action from above that generates rhombic designs. By the form of $D$ in (3.2), it holds for $\sigma^2(x)$ that

$$
\sigma^2(g(x)) = f(g(x))^T D f(g(x))
= f(x)^T Q_g^T D Q_g f(x).
$$

As $Q_g^T D Q_g = D$, it follows that $\sigma^2(x) = \sigma^2(g(x))$. \qed
The information matrix of a rhombic design is

\[ M(\xi) = \sum_{\ell=0}^{K} \frac{w_\ell}{N_\ell \sigma^2(x_\ell)} \sum_{x \in O_\ell} f(x)f(x)^T. \] (3.3)

To compute the matrix \( \sum_{x \in O_\ell} f(x)f(x)^T \), see that the information matrix is for each orbit structured into a scalar entry in the upper left corner and a \( K \times K \)-dimensional lower right completely symmetric block matrix. This leads to the following Lemma:

**Lemma 3.4.** In the setting of Section 3, a rhombic design \( \xi \) has an information matrix of the form

\[ M(\xi) = \begin{pmatrix} m_0(\xi) & 0 \\ 0 & M_1(\xi) \end{pmatrix}, \]

where \( m_0(\xi) = \sum_{\ell=0}^{K} \frac{w_\ell}{\sigma^2(x_\ell)} \) and \( M_1(\xi) = \sum_{\ell=0}^{K} \frac{w_\ell}{N_\ell \sigma^2(x_\ell)} \sum_{x \in O_\ell} xx^T = (m_1(\xi) - m_2(\xi))I_K + m_2(\xi)1_K1_K^T \) is a completely symmetric \( K \times K \)-dimensional matrix.

**Proof.** As for every \( x \in [-1, 1]^K \), \(-x\) lies in the same orbit as \( x \). Therefore, as

\[ f(x)f(x)^T + f(-x)f(-x)^T = 2 \begin{pmatrix} 1 & 0 \\ 0 & xx^T \end{pmatrix}, \]

\( M(\xi) \) is of the form

\[ M(\xi) = \begin{pmatrix} m_0(\xi) & 0 \\ 0 & M_1(\xi) \end{pmatrix}, \]

where

\[ M_1(\xi) = \sum_{\ell=0}^{K} \frac{w_\ell}{N_\ell \sigma^2(x_\ell)} \sum_{x \in O_\ell} xx^T \]

is a \( K \times K \)-dimensional symmetric matrix and

\[ m_0(\xi) = \sum_{\ell=0}^{K} \frac{w_\ell}{\sigma^2(x_\ell)}. \]

Now, \( \sum_{x \in O_\ell} xx^T \) is completely symmetric, because of the permutation invariance of the orbit \( O_\ell \). As the (weighted) sum of completely symmetric matrices is completely symmetric itself, the Lemma follows. \( \square \)

We defined \( \Gamma(\xi) \) as the matrix in the quadratic form in (2.2) coming from the equivalence theorem. By Lemma 3.4, denoting the lower right \( K \times K \)-submatrix of \( D \) by \( D_1 \),

\[ \Gamma(\xi) = \begin{pmatrix} (K+1)d_0 - m_0(\xi)^{-1} & 0 \\ 0 & (K+1)D_1 - M_1(\xi)^{-1} \end{pmatrix}. \]

\( \Gamma(\xi) \) has the same block structure as \( D \) and \( M(\xi) \) with completely symmetric lower right block of dimension \( K \times K \), so

\[ \Gamma(\xi) = \begin{pmatrix} \gamma_0 & 0 \\ 0 & (\gamma_1 - \gamma_2)I_K + \gamma_21_K1_K^T \end{pmatrix} \]

with \( \gamma_0 = (K+1)d_0 - m_0(\xi)^{-1} \).

We remind the reader of the following well-known fact:

The inverse of a completely symmetric \( K \times K \) matrix \( A \) with

\[ A = (a_1 - a_2)I_K + a_21_K1_K^T \]
\[ A^{-1} = \frac{1}{a_1 - a_2} I_K - \frac{a_2}{(a_1 - a_2)(a_1 + (K-1)a_2)} K^{-1} T_K. \]

With these preparatory results, we are able to investigate the optimality of rhombic designs.

4. Rhombic Designs and the Equivalence Theorem

4.1. Rhombic Vertex Designs. This section studies rhombic designs with all design points on the vertices of the hypercube. We will use the Kiefer-Wolfowitz equivalence theorem to investigate how the D-optimality of a rhombic vertex design depends on \( D \). The investigation leads to the following theorem:

**Theorem 4.1.** For the given model 3.1 with a dispersion matrix as defined in 3.2, let \( \xi^* \) be a rhombic vertex design. If either

1. \( K \) is even or
2. \( K \) is odd and \( \xi^* \) has support only on \( O \),

then \( \xi^* \) is D-optimal if and only if the matrix \( \Gamma(\xi^*) = pD - M(\xi^*)^{-1} \) is a diagonal matrix

\[ \Gamma(\xi^*) = \begin{pmatrix} \gamma_0 & 0 \\ 0 & \gamma_I K \end{pmatrix}, \]

with \( \gamma_0 \geq 0 \) and \( \gamma_1 = -\frac{\gamma_0}{N} \).

**Proof.** \(" \Rightarrow \)" Assuming \( \gamma_0 \geq 0 \), \( \gamma_1 = -\frac{\gamma_0}{N} \) and \( \gamma_2 = 0 \), it follows that

\[ \psi(x; \xi^*) = f(x)^T \Gamma(\xi^*) f(x) \]

\[ = \gamma_0 + \gamma_1 ||x||^2 \]

\[ = \gamma_0 \left( 1 - \frac{||x||^2}{K} \right) \]

Hence \( \psi(x; \xi^*) \geq 0 \) for all \( x \in [-1, 1]^K \), as \( ||x||^2 \leq K \). Therefore, the D-optimality follows from Theorem 2.1.

" \Rightarrow \)" According to (2.2) in Theorem 2.1, a design \( \xi^* \) is D-optimal if and only if

\[ (4.1) \]

\[ \psi(x; \xi^*) = f(x)^T \Gamma(\xi^*) f(x) \geq 0 \]

for all \( x \in X \). Furthermore, by Corollary 2.2, we know that

\[ (4.2) \]

\[ \psi(x; \xi^*) = 0 \]

for all support points of \( \xi^* \). Now, by (4.2), it follows for any design point \( x_\ell \in O_\ell(1) \), that

\[ (4.3) \]

\[ \psi(x_\ell, \xi^*) = \gamma_0 + K\gamma_1 + (K(K-1) - 4\ell(K-\ell))\gamma_2 = 0. \]

Assuming that the design is supported on at least two different orbits, this directly implies that \( \gamma_2 \) equals zero, as \( (K(K-1) - 4\ell(K-\ell)) \) is strictly monotone for \( 0 \leq \ell \leq K \).

Now, say that \( \xi^* \) is only supported on a single orbit \( O_\ell \) with \( 0 \leq \ell \leq K \) for even \( K \) and \( 0 \leq \ell < K \) for odd \( K \). If \( \ell = 0 \), it is easy to see that the information matrix is singular, therefore such a design cannot be D-optimal. The same is true for even \( K \) when \( \ell = K \) using the same argument as in [3]. It holds that \( \psi(x_\ell, \xi^*) = 0 \) for all \( x_\ell \in O_\ell(1) \). From the D-optimality of \( \xi^* \) it follows that \( \psi(x_{\ell-1}, \xi^*) \geq 0 \) for \( x_{\ell-1} \in O_{\ell-1}(1) \) and \( \psi(x_{\ell+1}, \xi^*) \geq 0 \) for \( x_{\ell+1} \in O_{\ell+1}(1) \), so in the orbits with one less or one more negative entry in the vector. If \( \gamma_2 \neq 0 \), this would imply \( \psi(x_{\ell-1}, \xi^*) > \psi(x_\ell, \xi^*) \) or \( \psi(x_{\ell+1}, \xi^*) > \psi(x_\ell, \xi^*) \). This contradicts the assumed D-optimality of \( \xi^* \), therefore \( \gamma_2 = 0 \) holds.
It follows from (4.3) and $\gamma_2 = 0$ that

$$\gamma_1 = -\frac{\gamma_0}{K}.$$ 

This implies that

$$\psi(x; \xi^*) = \gamma_0 \left(1 - \frac{||x||^2}{K}\right).$$

and therefore $\gamma_0 \geq 0$ when the design on the vertices is $D$-optimal.

Note that $\psi(x, \xi^*) = 0$ for all vertices $x$ of the hypercube as for those we have $||x||^2 = K$.

**Corollary 4.2.** A rhombic vertex design with support on either at least two orbits or an orbit $O_\ell(1)$ with $1 \leq \ell < K$ can only be $D$-optimal when

$$(d_1 - d_2)(d_1 + (K - 1)d_2) - d_0(d_1 + (K - 2)d_2) \leq 0$$

**Proof.** From Theorem 4.1 we obtain that

$$\gamma_0 \geq 0, \ \gamma_1 = -\frac{\gamma_0}{K}, \ \gamma_2 = 0$$

is equivalent to the $D$-optimality of a rhombic vertex design with support on at least two orbits or an orbit $O_\ell(1)$ with $1 \leq \ell < K$. The equation system $\{\gamma_1 = -\frac{\gamma_0}{K}, \gamma_2 = 0\}$ has two solutions $m_0(\xi)^\pm$ for $m_0(\xi)$ in dependence of $d_0, d_1, d_2$ and $p$ that we obtained with MATHEMATICA:

$$m_0^\pm(\xi) = \frac{d_0(p + 1) + (p - 1) (d_1 p + d_1 + d_2 p^2 - 3 d_2 p)}{2 p (d_1^2 + d_0(p - 1)(2 d_1 + d_2(p - 3)) + (p - 1)^2 (d_1 - d_2)(d_1 + d_2(p - 2)))}$$

$$\pm \frac{(p - 1) \sqrt{d_0^2 + 2 d_0(d_1(p - 1) + d_2(p - 3)p) + (p - 1) (d_1^2(p - 1) + 2 d_1 d_2(p - 3)p + d_2^2 p(p^2 - 5p + 8))}}{2 p (d_1^2 + d_0(p - 1)(2 d_1 + d_2(p - 3)) + (p - 1)^2 (d_1 - d_2)(d_1 + d_2(p - 2)))}.$$ 

Now, with $\gamma_0 = pd_0 - \frac{1}{m_0(\xi)}$ we see that

$$\Leftrightarrow 0 \geq \left(pd_0 - \frac{1}{m_0(\xi)}\right) \left(pd_0 - \frac{1}{m_0(\xi)}\right)$$

$$\Leftrightarrow 0 \geq p^2 d_0^2 m_0^+(\xi) m_0^-(\xi) - (m_0^+(\xi) + m_0^-(\xi)) + 1$$

$$\Leftrightarrow 0 \geq (d_1 - d_2)(d_1 + (K - 1)d_2) - d_0(d_1 + (K - 2)d_2).$$

To derive the corollary, check that $pd_0 - \frac{1}{m_0(\xi)}$ is always negative on $C_K$, so that

$$0 \leq pd_0 - \frac{1}{m_0(\xi)}$$

$$\Leftrightarrow 0 \leq pd_0 - \frac{1}{m_0^+(\xi)}$$

$$\Leftrightarrow 0 \geq \left(pd_0 - \frac{1}{m_0^+(\xi)}\right) \left(pd_0 - \frac{1}{m_0^+(\xi)}\right)$$

on $C_K$. This implies the corollary.

**Remark 4.3.** Theorem 4.1 gives a semi-algebraic description of the optimality region of rhombic vertex design in the design weights and the coefficients $d_0, d_1$ and $d_2$. This means that

$$\left\{\gamma_0 \geq 0, \gamma_1 = -\frac{\gamma_0}{K}, \gamma_2 = 0\right\}$$
can be interpreted as a semi-algebraic set if one takes into account the constraints that $\xi$ is a design and $(d_0, d_1, d_2)^T \in C_K$. This allows us to obtain symbolic solutions for the design weights $\xi^*_\ell := \xi^*(O_\ell(1))$ in dependence of the coefficients $d_0, d_1$ and $d_2$.

4.2. Non-vertex (rhombic) designs. Instead of restricting to rhombic designs we will discuss a broader class of designs, namely designs with a design point in the interior of the hypercube. This design class naturally includes non-vertex rhombic designs.

**Theorem 4.4.** A design $\xi^*$ with at least one design point in the interior of the hypercube is $D$-optimal if and only if $M(\xi^*) = \frac{1}{K+1} D^{-1}$, which means that $\Gamma(\xi^*) = (K+1) D - M(\xi^*)^{-1}$ is zero.

**Proof.** $\Leftarrow$ $M(\xi^*) = \frac{1}{K+1} D^{-1}$ implies the $D$-optimality of $\xi^*$ by Theorem 2.1.

$\Rightarrow$ By Theorem 2.1 and Corollary 2.2, $\xi^*$ is $D$-optimal if and only if $\psi(x; \xi^*) = f(x)^T ((K+1) D - M(\xi^*)^{-1}) f(x) \geq 0$

for all $x \in [-1, 1]^K$ and

$$\psi(x; \xi^*) = 0,$$

if $x$ is a design point of $\xi^*$. Now, with $\Gamma(\xi^*) = (K+1) D - M(\xi^*)^{-1}$,

$$\psi(x; \xi^*) = f(x)^T \Gamma(\xi^*) f(x) = \gamma_0 + \gamma_1 ||x||^2 + 2\gamma_2 \sum_{1 \leq \ell < j \leq K} x_\ell x_j$$

is a quadratic polynomial. It holds that

(4.4) $$\gamma_0 + \gamma_1 ||x||^2 + 2\gamma_2 \sum_{1 \leq \ell < j \leq K} x_\ell x_j = 0$$

for design points $x$ of $\xi^*$. As $\xi^*$ is a non-vertex design there is an interior point $x_1$ and additionally $K$ further design points $x_2, ..., x_{K+1}$ such that $x_1, ..., x_{K+1}$ span $\mathbb{R}^K$ because the information matrix $M(\xi^*)$ needs to be non-singular. Fix the affine subspace generated by $x_1$ and $x_2$. It holds that $\psi(x_1, \xi^*) = \psi(x_2, \xi^*) = 0$ and $\psi(\lambda x_1 + (1-\lambda)x_2, \xi^*) \geq 0$ for all $\lambda \in [0, 1]$ and additionally for some $\lambda < 0$ because $x_1$ is an interior point of $[-1, 1]^K$. On the affine subspace generated by $x_1$ and $x_2$, $\psi(x; \xi^*)$ is a quadratic polynomial in $\lambda$. The only quadratic polynomial that is zero on at least two points and non-negative on at least one point on the line segment between these points as well as for at least one point on the line outside this segment is the zero polynomial. Recursively, this can be extended to higher dimensions. Therefore,

$$f(x)^T ((K+1) D - M(\xi^*)^{-1}) f(x) \geq 0$$

for all $x \in [-1, 1]^K$ can only be achieved as an equation, so

$$(K+1) D - M(\xi^*)^{-1} = 0.$$ 

Hence, the Theorem follows. \hfill \Box

**Corollary 4.5.** An invariant design $\xi^*$ with a design point in the interior of the hypercube can only be $D$-optimal if the first diagonal entry of $D^{-1}$ is larger than the second, which means that

$$(d_1 - d_2)(d_1 + (K-1)d_2) - d_0(d_1 + (K-2)d_2) > 0.$$ 

**Proof.** According to Theorem 4.4, an invariant design $\xi^*$ with a design point in the interior of the hypercube is $D$-optimal if and only if $\Gamma(\xi^*) = 0$. Now, this implies that $M(\xi^*) = \frac{1}{K+1} D^{-1}$ and therefore

$$m_0(\xi^*) = \frac{1}{(K+1)d_0}, \quad m_1(\xi^*) = \frac{d_1 + (K-2)d_2}{(K+1)(d_1 - d_2)(d_1 + (K-1)d_2)},$$

and therefore...
where \( m_1(\xi^*) \) denotes the diagonal entries of \( M_1(\xi^*) \). It is easy to see that \( m_0(\xi^*) > m_1(\xi^*) \) for all designs with interior design points, so we obtain

\[
m_0(\xi^*) > m_1(\xi^*)
\]

\[
\iff (d_1 - d_2)(d_1 + (K - 1)d_2) - d_0(d_1 + (K - 2)d_2) > 0.
\]

**Remark 4.6.** Theorem 4.4 describes the semi-algebraic structure of the optimality area of non-vertex designs. We see that this structure is given by the non-negative real part of an algebraic variety, so the vanishing set of a collection of polynomials under the constraints of the model cone \( C_K \) and the design simplex. This means that we can obtain symbolic solutions for the optimal designs weights and design points in dependence of the coefficients \( d_0, d_1 \) and \( d_2 \) by studying the set \( \{ \Gamma(\xi) = 0 \} \) under the imposed constraints.

5. **Rhombic Designs for \( K \in \{2, 3, 4, 5\} \)**

This section investigates for which values \( d_0, d_1 \) and \( d_2 \) we find a vertex or non-vertex rhombic design for \( 2 \leq K \leq 5 \).

5.1. **The case \( K = 2 \)**. The results of this section where first calculated by hand and later confirmed with a Mathematica implementation of Theorems 4.1 and 4.4. For \( K = 2 \), it is

\[
(5.1) \quad D = \begin{pmatrix} d_0 & 0 & 0 \\ 0 & d_1 & d_2 \\ 0 & d_2 & d_1 \end{pmatrix}
\]

where \( |d_2| < d_1 \) and the variance of each design point is equal to

\[
\sigma^2(x) = f(x)^T D f(x) = d_0 + d_1(x_0^2 + x_1^2) + 2d_2x_0x_1.
\]

The symmetric properties of the covariance structure with respect to the random effects of the two attributes, \( \text{Var}(b_1) = \text{Var}(b_2) = d_1 \), motivates us to consider as candidates for the \( D \)-optimal designs the following rhombic designs \( \xi_{x_0 \circ x_1, w} \) consisting of the four design points \((x_0, x_0), (-x_0, -x_0), (-x_1, x_1) \) and \((x_1, -x_1)\) for \( x_0, x_1 \in [-1, 1] \) which form a centered rhombus within the design region. Therefore we have \( \mathcal{O}_0(x_0) = \{(x_0, x_0), (-x_0, -x_0)\} \) and \( \mathcal{O}_1(x_1) = \{(-x_1, x_1), (x_1, -x_1)\} \), so that \( \sigma^2(x_0, x_0) = \sigma^2(-x_0, -x_0) \) and \( \sigma^2(-x_1, x_1) = \sigma^2(x_1, -x_1) \). It follows that we can deal with two distinct weights \( w_0 = \xi(\mathcal{O}_0(x_0)) \) and \( w_1 = \xi(\mathcal{O}_1(x_1)) \) since the sum of the weights of all orbits is equal to 1 we can set \( w_0 = w \) and \( w_1 = 1 - w \). The information matrix for \( \xi_{x_0 \circ x_1, w} \) results in

\[
M(\xi_{x_0 \circ x_1, w}) = \begin{pmatrix}
\frac{w}{\sigma^2(x_0, x_0)} + \frac{1-w}{\sigma^2(-x_1, x_1)} & 0 & 0 \\
0 & \frac{w}{\sigma^2(x_0, x_0)} - \frac{1-w}{\sigma^2(-x_1, x_1)} & \frac{1-w}{\sigma^2(-x_1, x_1)} \\
0 & \frac{1-w}{\sigma^2(x_0, x_0)} & \frac{w}{\sigma^2(x_0, x_0)} + \frac{1-w}{\sigma^2(-x_1, x_1)}
\end{pmatrix}
\]

with determinant

\[
\det(M(\xi_{x_0 \circ x_1, w})) = 4 \frac{w \sigma^2(-x_1, x_1) + (1-w) \sigma^2(x_0, x_0)}{\sigma^2(x_0, x_0) \sigma^2(-x_1, x_1)} w (1-w) x_0^2 x_1^2.
\]

Maximizing the determinant with respect to the variables \( x_0, x_1 \) and \( w \) leads to the following results:

Note that \( d_2 = 0 \) is excluded as this was already settled in [5].

**Theorem 5.1.** In the heteroscedastic model of two-factorial multiple regression on \([-1, 1]^2 \) with dispersion matrix (5.1) it follows:
(i) If $d_0 \leq d_1 - |d_2|$, the design $\xi_{x_0^*, x_1^*, 0.5}$ is $D$-optimal with

$$x_0^* = \sqrt{\frac{d_0}{d_1 + d_2}}, \quad x_1^* = \sqrt{\frac{d_0}{d_1 - d_2}}.$$ 

(ii) If $d_1 - |d_2| \leq d_0 \leq \frac{d_1^2 - d_2^2}{d_1}$ the design $\xi_{x_0^*, x_1^*, w^*}$ is $D$-optimal with

$$w^* = \frac{2}{3} - \frac{d_0}{6(d_1 - d_2)}, \quad x_0^* = \sqrt{\frac{d_1 - d_2}{d_1 + d_2} \frac{d_0}{2(d_1 - d_2) - d_0}}, \quad x_1^* = 1, \quad \text{if } d_2 > 0,$$

$$w^* = \frac{1}{3} + \frac{d_0}{6(d_1 + d_2)}, \quad x_0^* = 1, \quad x_1^* = \sqrt{\frac{d_1 + d_2}{d_1 - d_2} \frac{d_0}{2(d_1 + d_2) - d_0}}, \quad \text{if } d_2 < 0.$$ 

(iii) If $\frac{d_1^2 - d_2^2}{d_1} \leq d_0$ the vertex design $\xi_{1, 1, w^*}$ is $D$-optimal where $w^*$ solves the equation

$$2 \left( d_2 \left( 6w^2 - 6w + 1 \right) + d_1(1 - 2w) \right) + d_0(1 - 2w) = 0.$$ 

Proof. Check with Theorems 4.1 and 4.4, that the designs are optimal. \qed

(i) and (ii) describe non-vertex rhombic designs, while (iii) describes the rhombic design with support on the vertices of the square. The Theorem shows that there is a $D$-optimal rhombic design for all $(d_0, d_1, d_2)^T \in C_2$. Figures 3 and 4 visualize the different optimality regions in Theorem 5.1 for $d_0 = 1$. Note that the region only depends on the quotients $\frac{d_1}{d_0}$ and $\frac{d_2}{d_0}$, so the choice of $d_0$ is arbitrary.

5.2. The case $K = 3$. The following Theorem results from a MATHEMATICA implementation of Theorems 4.1 and 4.4.
Figure 4. Assembling the parameter regions of rhombic designs for $K = 2$

**Theorem 5.2.** For the setting from Section 3 with $K = 3$, so

$$D = \begin{pmatrix} d_0 & 0 & 0 & 0 \\ 0 & d_1 & d_2 & d_2 \\ 0 & d_2 & d_1 & d_2 \\ 0 & d_2 & d_2 & d_1 \end{pmatrix},$$

where $-d_2^2 < d_2 < d_1$ it follows:

(i) If either $d_0 < d_3 + d_2 \land 0 < d_2 < \frac{d_1}{2}$ or $d_0 < \frac{(d_1 + 2d_2)^2}{d_1 - 2d_2} \land d_2 < 0$, the design with $w^* = \frac{1}{4}$ and

$$x_0^* = \sqrt{\frac{d_0(d_1 - 2d_2)}{d_1 + 2d_2}}, \quad x_1^* = \sqrt{\frac{d_0}{d_1 - 2d_2}},$$

is $D$-optimal.

(ii) If $d_2 < \frac{d_1}{2} \land d_0 < \frac{(d_1 - d_2)(d_1 + 2d_2)}{d_1 + d_2}$, it holds that the design with

$$w^* = \frac{3d_0 - 7d_1 + 10d_2}{-16(d_1 - d_2)}, \quad x_0^* = \sqrt{\frac{d_0(-d_1 + 2d_2)}{(d_1 + 2d_2)(3d_0 - 4d_1 + 4d_2)}}, \quad x_1^* = 1,$$

is $D$-optimal.

(iii) If $d_2 < \frac{d_1}{2} \land d_0 < \frac{(d_1 - d_2)(d_1 + 2d_2)}{d_1 + d_2}$, it holds that the design with

$$x_0^* = 1, \quad x_1^* = \sqrt{\frac{3d_0(d_1 + 2d_2)}{2d_0d_2 - d_0d_1 - 8d_2^2 + 4d_2d_1 + 4d_1^2}},$$

$$w^* = \frac{(d_1 - 2d_2)(d_0 + 3d_1 + 6d_2)}{16(d_1 - d_2)(d_1 + 2d_2)},$$

is $D$-optimal.
Figure 5. Parameter regions for $K = 3$: Figure (A) shows parameter regions of rhombic non-vertex designs, while Figure (B) shows parameter regions of rhombic vertex designs.

(iv) If $d_2 \neq 0$ and $(d_0(d_1 + d_2) \geq (d_1 - d_2)(d_1 + 2d_2) \land d_2 \leq \frac{d_1}{2}) \lor (\frac{d_1}{2} < d_2 \land 3d_0 + 9d_1 > 22d_2)$, then the design with $x_0 = x_1 = 1$ and
\[
 w^* = \frac{3d_0^2 + 22d_0d_2 + 18d_0d_1 - 120d_2^2 + 66d_2d_1 + 27d_1^2}{64d_2(d_0 - 3d_2 + 3d_1)} \\
 - \frac{3\sqrt{(d_0 - 2d_2 + 3d_1)^2(d_0^2 + 8d_0d_2 + 6d_0d_1 + 48d_2^2 + 24d_2d_1 + 9d_1^2)}}{64d_2(d_0 - 3d_2 + 3d_1)}
\]
is $D$-optimal.

Proof. For the cases (i), (ii), (iii) check that the equation $\frac{1}{4}D^{-1} = M(\xi^*)$ holds and the model constraints are satisfied. For the fourth case, check that $m_0(\xi^*) \geq \frac{1}{4d_0^2}$ and that the model constraints are satisfied. □

Note that not all settings of $(d_0, d_1, d_2)$ are covered by Theorem 5.2 and that the described design areas are not disjoint. (ii) and (iii) describe the same optimality area that also contain area (i). Figures 5 and 6 show the optimality area for $d_0 = 1$ in the $(d_1, d_2)$-space. Again, the region only depends on the quotients $\frac{d_1}{d_0}$ and $\frac{d_2}{d_0}$, so the choice of $d_0$ is arbitrary. The area where we did not find an optimal rhombic design is given by $\frac{d_1}{2} < d_2 \land 3d_0 + 9d_1 \leq 22d_2$.

5.3. The cases $K = 4$ and $K = 5$. For $K = 4$ and $K = 5$ there are up to three orbits for rhombic designs. To compute an optimal rhombic vertex design, we let $O_\ell(x_\ell)$ denote the orbits of rhombic design points and choose $x_0 = x_1 = x_2 = 1$, such that the weights are $w_\ell = \xi(O_\ell(1))$ and check the conditions in Theorem 4.1 for optimality. The different optimality areas are shown in Figure 7. Again, in the red region, an rhombic design with interior points is $D$-optimal, while in the yellow area, a rhombic vertex design is $D$-optimal. The separating line is again given by the equality of the first and the second diagonal entry of $D^{-1}$, see Corollary 4.2 and Corollary 4.5. We see a similar
Figure 6. Assembling the parameter regions for $K = 3$

Figure 7. Assembling the optimality regions for $K = 4$ and $K = 5$

structure as for $K = 2$ and $K = 3$. For $K = 4$, there is a $D$-optimal rhombic design for every point in $C_4$, while for $K = 5$, in the region above $d_2 = \frac{d_1}{2}$ there is only a small area where rhombic designs are $D$-optimal, similar to the case for $K = 3$.

Remark 5.3. The optimality regions shown in the figures for $K \in \{2, 3, 4, 5\}$ are given in the $(d_1, d_2)$-space while $d_0 = 1$. As before, the region only depends on the quotients $\frac{d_1}{d_0}$ and $\frac{d_2}{d_0}$, so the choice of $d_0 = 1$ is arbitrary. $D$-optimal designs and the corresponding parameter regions where they are optimal can be found by studying the semi-algebraic sets as described in Remark 4.6 and Remark 4.3. A convenient way to generate the images showing the optimality regions is therefore to use the
Resolve and RegionPlot commands of Mathematica to compute and plot these regions. This was done for $K \in \{2, 3, 4, 5\}$.

6. Discussion

In the preceding sections optimality regions have been investigated for certain invariant designs in a multiple linear regression model on the hypercube with invariant correlation structure of the random coefficients. It has been shown that for the introduced class of rhombic designs, it is possible to decide whether a $D$-optimal design is either supported on the vertices of the hypercube or has interior design points by evaluating a quadratic polynomial depending on the covariance matrix of the random coefficients. This result relies on the Kiefer-Wolfowitz equivalence theorem. The equation separating the two optimality regions is given as the equality of the diagonal entries of $D^{-1}$.

The results of Theorem 4.4 hold not only for rhombic designs but for all designs with an interior design point, independently of invariance considerations. This means that the $D$-optimality of designs with interior points is equivalent to the equation $M(\xi^*) = \frac{1}{p} D^{-1}$.

An important observation is the apparent non-existence for $D$-optimal rhombic designs for certain values of the entries $D$. For small dimensions, we have observed that for even $K$, we could always find a $D$-optimal rhombic design for any $D$, while this has not been true for odd $K$. With respect to our findings, we conjecture the following:

**Conjecture 6.1.** For even $K$, there is a $D$-optimal rhombic design for all $(d_0, d_1, d_2)^T \in \mathcal{C}_K$. For odd $K$, there is a $D$-optimal rhombic design for all $(d_0, d_1, d_2)^T \in \mathcal{C}_K$ with $d_2 \leq \frac{d_1}{2}$.

**Acknowledgements.** Work supported by grants HO 1286/6, SCHW 531/15 and 314838170, GRK 2297 MathCoRe of the Deutsche Forschungsgemeinschaft DFG.

**References**

[1] A. C. Atkinson, V. V. Fedorov, A. M. Herzberg, and R. Zhang. Elemental information matrices and optimal experimental design for generalized regression models. *Journal of Statistical Planning and Inference*, 144:81 – 91, 2014. International Conference on Design of Experiments.

[2] V. V. Fedorov. *Theory of optimal experiments*. Academic Press, New York-London, 1972. Translated from the Russian and edited by W. J. Studden and E. M. Klimko, Probability and Mathematical Statistics, No. 12.

[3] F. Freise, H. Holling, and R. Schwabe. Optimal designs for two-level main effects models on a restricted design region. *Journal of Statistical Planning and Inference*, 204:45 – 54, 2020.

[4] P. A. Freund and H. Holling. Creativity in the classroom: A multilevel analysis investigating the impact of creativity and reasoning ability on gpa. *Creativity Research Journal - CREATIVITY RES J*, 20:309–318, 08 2008.

[5] U. Graßhoff, A. Doebler, H. Holling, and R. Schwabe. Optimal design for linear regression models in the presence of heteroscedasticity caused by random coefficients. *J. Statist. Plann. Inference*, 142(5):1108–1113, 2012.

[6] U. Graßhoff, H. Holling, and R. Schwabe. On optimal design for a heteroscedastic model arising from random coefficients. *Proceedings of the 6th St. Petersburg Workshop on Simulation*, 01 2009.

[7] J. Kiefer. General equivalence theory for optimum designs (approximate theory). *The Annals of Statistics*, 2(5):849–879, 1974.

[8] M. Patan and B. Bogacka. Efficient sampling windows for parameter estimation in mixed effects models. In *mODa 8—Advances in model-oriented design and analysis*, Contrib. Statist., pages 147–155. Physica-Verlag/Springer, Heidelberg, 2007.

[9] S.D. Silvey. *Optimal design: an introduction to the theory for parameter estimation*. Monographs on applied probability and statistics. Chapman and Hall, 1980.
