Optimal Symmetric Designs for a One-way Layout with Covariates

Rainer Schwabe*
Technische Hochschule Darmstadt, Fachbereich Mathematik,
Schloßgartenstraße 7, 64 289 Darmstadt, Germany

ABSTRACT: For the general class of Φ_q-criteria optimal desgins are characterized which reflect the inherent symmetry in a one-way layout with covariates. In particular, the eigenvalues of the covariance matrices are related to those in suitably chosen marginal models depending on the underlying interaction structure.

AMS 1991 classification: 62K05

Keywords: D-optimal design, A-optimal design, E-optimal design, symmetric desgin, symmetric model, marginal model, additive model, partial interactions, qualitative and quantitative factors.

1. Introduction.

In most experimental situations several factors are active which influence the outcome of the experiment. If the factors may, additionally, interact with each others it is usually hard to find a good, or even optimal, design.

In the case that one of the factors is qualitative its influence can be described by a one-way layout with a finite number of different treatments. In this situation it is reasonable to assume a further structure in the underlying linear model which is left unchanged by relabeling the levels of that qualitative factor. Essentially this is true if the structure of interactions does not depend on the actual level. Those models will be specified in Section 2. A good design should reflect this inherent symmetry property.

Optimal designs aim at miniminzing certain characteristics of the covariance matrix for the estimators of the unknown parameters in the model. E. g. the common A-, D- and E-criteria are all based on the eigenvalues of the covariance matrix. In this paper we consider the larger family of Φ_q-criteria, and we optimize within the class of generalized designs (for more background informations on this topic we refer

*Research partly supported by grant Ku 719/2-2 of the Deutsche Forschungsgemeinschaft

to Kiefer, 1974, or the mongraphs by Bandemer et al., 1977, Atkinson and Donev, 1992, and Pukelsheim, 1993, besides others).

Φ_q-criteria are invariant with respect ot permutations of the levels of the qualitative factor. Hence, optimal designs can be found in the essentially complete class of symmetric (invariant) designs. Kurotschka, Schwabe and Wierich (1992) derived D-optimal designs for certain interaction structures in linear models with both qualitative and quantitative factors of influence. This approach was extended by Schwabe (1996, Section 6.2) to general models symmetric in one factor and to the A-criterion. Recently, for those classes treated by Kurotschka et al. (1992) E-optimal designs were calculated by Schwabe (1996b). (See Schwabe, 1996, for further references.)

Besides the general symmetric model also the related marginal models are introduced in the next section. Then, in Section 3, results are obtained for the class of all Φ_q-criteria which aim at minimizing the "q-norm" of the eigenvalues of the covariance matrix. The proofs are based on an explicit representation of the eigenvalues in terms of the marginal models. Hence, the optimal designs are characterized by conditions on their marginals which leads to a substantial reduction of the complexity of the optimization problem.

Related results are obtained in the subsequent sections 4 and 5 for invariant parts of the parameter vectors and for models with an explicitly included mean function. In particular, if the only interest is in the overall mean function then the symmetrization leads to a uniform improvement. Finally, extensions to higher dimensional models are indicated in Section 6.

2. Symmetric models.

In this paper we consider the two-factor linear models

$$E(Y(j,x)) = f_0(x)^\top \beta_0 + f_1(x)^\top \beta_j, \tag{1}$$

$j \in 1, ..., J,\ x \in \mathcal{X}$, which show an invariant structure with respect to permutations of the levels j of the qualitative factor. The factor x can be of arbitrary type: quantitative, qualitative, or higher dimensional itself.

To be more precise, let $\beta = (\beta_0^\top, \beta_1^\top, ..., \beta_J^\top)^\top$ be the vector of unknown parameters in (1) and $a_1 = (\mathbf{1}_{\{1\}}, ..., \mathbf{1}_{\{J\}})^\top$ the vector of indicator functions $\mathbf{1}_{\{j\}}$ on $\{1, ..., J\}$. The model (1) can be rewritten in the common general linear model notation $E(Y) = a^\top \beta$ as

$$E(Y(j,x)) = (f_0(x)^\top,\ a_1(j)^\top \otimes f_1(x)^\top)\beta, \tag{2}$$

where "$\otimes$" denotes the usual Kronecker (tensor) product of matrices and vectors, respectively. It can easily be verified that permutations of the levels j of the qualitative factor induce orthogonal transformations of the regression function $a = (f_0^\top, a_1^\top \otimes f_1^\top)^\top$ in the representation (2) (see Schwabe, 1996, p 83).

For every criterion which is convex and invariant with respect to these transformations any design δ is dominated by its symmetrization $\bar{\delta}_1 \otimes \delta_2$ where $\bar{\delta}_1$ is the uniform design on $\{1, ..., J\}$, i.e. $\bar{\delta}_1(j) = \frac{1}{J}$, δ_2 is the second marginal of δ, i.e. $\delta_2(x) = \sum_{j=1}^{J} \delta(j, x)$, and "$\otimes$" denotes the product of designs, i.e. $\delta_1 \otimes \delta_2(j, x) = \delta_1(j)\delta_2(x)$. The symmetric product designs $\bar{\delta}_1 \otimes \delta_2$ constitute an essentially complete class. It remains to optimize with respect to the second marginal design δ_2.

To characterize the optimal marginal design δ_2^* particular marginal models are of interest in which only the second factor x and its corresponding regression function f_0 and f_1 are involved. The standard marginal model

$$E(Y_2(x)) = f_0(x)^\top \beta_0 + f_1(x)^\top \beta_1, \tag{3}$$

$x \in \mathcal{X}$, is obtained by fixing the first factor to some arbitrary level. Note that the specific choice of the first level does not affect the struture of (3), i.e. the regression functions do not depend on j. To avoid technicalities we assume throughout the paper that the components of $f = (f_0^\top, f_1^\top)^\top$ are linearly independent on $\mathcal{X}$. This guarantees that the whole parameter vector can be estimated in both the marginal model (3) and the two-factor model (1) if the designs are rich enough. In particular, the information matrix $\mathbf{I}(\bar{\delta}_1 \otimes \delta_2) = \int a a^\top d(\bar{\delta}_1 \otimes \delta_2)$ is regular if and only if the corresponding marginal information matrix $\mathbf{I}_2(\delta_2) = \int f f^\top d\delta_2$ is regular. A second important model is the weighted marginal model

$$E(\tilde{Y}_2(x)) = f_0(x)^\top \beta_0 + \frac{1}{\sqrt{J}} f_1(x)^\top \beta_1, \tag{4}$$

$x \in \mathcal{X}$. This is obtained from the standard marginal model (3) by a reparametrization which increases the influence of β_1 by a factor related to the number J of levels of the first factor. Finally, we are also interested in the restricted marginal model

$$E(Y_2^{(1)}(x)) = f_1(x)^\top \beta_1, \tag{5}$$

$x \in \mathcal{X}$, which consists of that part of the regression function interacting with the first factor. The corresponding information matrices are denoted by

$$\tilde{\mathbf{I}}_2(\delta_2) = \begin{pmatrix} \int f_0 f_0^\top d\delta_2 & \frac{1}{\sqrt{J}} \int f_0 f_1^\top d\delta_2 \\ \frac{1}{\sqrt{J}} \int f_1 f_0^\top d\delta_2 & \frac{1}{J} \int f_1 f_1^\top d\delta_2 \end{pmatrix} \tag{6}$$

and $\mathbf{I}_2^{(1)}(\delta_2) = \int f_1 f_1^\top d\delta_2$ for the weighted marginal model (4) and the restricted marginal model (5), respectively.

3. Optimal designs.

We start with the inference on the whole parameter vector $\beta = (\beta_0^\top, \beta_1^\top, ..., \beta_J^\top)^\top$. The Φ-criteria aim at minimizing the "q-norm" $\Phi_q(\delta) = \sum_{i=1}^p \nu_i^q$ of the eigenvalues $\nu_1, ..., \nu_p$ of the covariance matrix $\mathbf{C}(\delta) = \mathbf{I}(\delta)^{-1}$, $0 < q < \infty$, or equivalently at maximizing the "inversed q-norm" $\Phi_q(\delta)^{-1} = (\sum_{i=1}^p \lambda_i^{-q})^{-1}$, of the eigenvalues $\lambda_1, ..., \lambda_p$ of the information matrix $\mathbf{I}(\delta)$. This includes the common A-criterion $\Phi_1(\delta) = \mathrm{trace}(\mathbf{C}(\delta))$, and the D- and E-criteria are obtained as limiting cases $\Phi_0(\delta) = \det(\mathbf{C}(\delta)) = \det(\mathbf{I}(\delta))^{-1}$ and $\Phi_\infty(\delta) = \lambda_{\max}(\mathbf{C}(\delta)) = \lambda_{\min}(\mathbf{I}(\delta))^{-1}$ respectively. These Φ_q-criteria are invariant under the permutations of the levels of the first factor because the eigenvalues of $\mathbf{I}(\delta) = \int a a^\top d\delta$ and, hence, of $\mathbf{C}(\delta)$ are preserved under the induced orthogonal transformations of the regression function a. Thus we can confine to the essentially complete class of symmetric designs $\bar{\delta}_1 \otimes \delta_2$. The corresponding information matrices are given by

$$\mathbf{I}(\bar{\delta}_1 \otimes \delta_2) = \begin{pmatrix} \int f_0 f_0^\top d\delta_2 & \frac{1}{J}\mathbf{1}_J^\top \otimes \int f_0 f_1^\top d\delta_2 \\ \frac{1}{J}\mathbf{1}_J \otimes \int f_1 f_0^\top d\delta_2 & \frac{1}{J}\mathbf{E}_J \otimes \int f_1 f_1^\top d\delta_2 \end{pmatrix} \tag{7}$$

where $\mathbf{E}_J$ and $\mathbf{1}_J$ denote the $J \times J$ identity matrix and the J-dimensional vector with all entries equal to one, respectively.

Lemma 1. *Let $\tilde{\lambda}_1, ..., \tilde{\lambda}_{p_2}$ be the eigenvalues of $\tilde{\mathbf{I}}_2(\delta_2)$ and let $\lambda_1^{(1)}, ..., \lambda_{p(1)}^{(1)}$ be the eigenvalues of $\mathbf{I}_2^{(1)}(\delta_2)$, respectively. Then the eigenvalues of $\mathbf{I}(\delta_1 \otimes \delta_2)$ are given by $\tilde{\lambda}_1, ..., \tilde{\lambda}_{p_2}$ and $J-1$ replicates of $\frac{1}{J}\lambda_1^{(1)}, ..., \frac{1}{J}\lambda_{p(1)}^{(1)}$.*

Proof. If $\tilde{z} = (\tilde{z}_0^\top, \tilde{z}_1^\top)^\top$ is an eigenvector of $\tilde{\mathbf{I}}_2(\delta_2)$ associated with the eigenvalue $\tilde{\lambda}_i$ partitioned according to the dimensions of f_0 and f_1, then $z = (\tilde{z}_0^\top, \frac{1}{\sqrt{J}}(\mathbf{1}_J \otimes \tilde{z}_1)^\top)^\top$ is an eigenvector of $\mathbf{I}(\bar{\delta}_1 \otimes \delta_2)$ with eigenvalue λ_i which proves the first part. For the remaining eigenvalues let $z_{(1)}$ be an eigenvector of $\mathbf{I}_2^{(1)}(\delta_2)$ associated with $\lambda_i^{(1)}$. There are $J-1$ linearly independent J-dimensional contrasts $\ell_1, ..., \ell_{J-1}$, i.e. $\ell_j^\top \mathbf{1}_J = 0$. Then $z = (\mathbf{0}, (\ell_j \otimes z_{(1)})^\top)^\top$ is an eigenvector of $\mathbf{I}(\bar{\delta}_1 \otimes \delta_2)$ with eigenvalue $\frac{1}{J}\lambda_i^{(1)}$. $\square$

With this preliminary result it is straightforward to characterize the optimal designs. To this end denote by $\Phi_{2,q}$, $\tilde{\Phi}_{2,q}$ and $\Phi_{2,q}^{(1)}$ the Φ_q-criterion functions in the marginal models (3) to (5), respectively.

Theorem 1. *The symmetric design $\bar{\delta}_1 \otimes \delta_2^*$ is*
(i) *Φ_q-optimal if δ_2^* minimizes $\tilde{\Phi}_{2,q}(\delta_2) + (J-1)J^q \Phi_{2,q}^{(1)}(\delta_2)$, for $0 < q < \infty$.*
(ii) *D-optimal if δ_2^* maximizes $\det(\mathbf{I}_2(\delta_2)) \det(\mathbf{I}_2^{(1)}(\delta_2))^{J-1}$.*
(iii) *E-optimal if δ_2^* is E-optimal in the weighted marginal model (4).*

Proof. By Lemma 1 we can calculate $\Phi_q(\bar\delta_1 \otimes \delta_2)$ from the marginals. For $0 < q < \infty$ we have $\Phi_q(\bar\delta_1 \otimes \delta_2) = \widetilde\Phi_{2,q}(\delta_2) + (J-1)J^q\Phi_{2,q}^{(1)}(\delta_2)$ which immediately establishes (i).

For the D-criterion $\det(\mathbf{I}(\bar\delta_1 \otimes \delta_2)) = J^{p_2+(J-1)p(1)}\det(\widetilde{\mathbf{I}}_2(\delta_2))\det(\mathbf{I}_2^{(1)}(\delta_2))^{J-1}$ and, additionally, $\det(\widetilde{\mathbf{I}}_2(\delta_2)) = c\det(\mathbf{I}_2(\delta_2))$, for some constant $c > 0$, as (4) is a reparametrization of (3) which, together, prove (ii).

Finally, for the E-criterion $\lambda_{\min}(\mathbf{I}(\bar\delta_1 \otimes \delta_2)) = \min(\lambda_{\min}(\widetilde{\mathbf{I}}_2(\delta_2)), \frac{1}{J}\lambda_{\min}(\mathbf{I}_2^{(1)}(\delta_2)))$. As the restricted marginal model (5) can be considered as a submodel of (4) after scaling by $\frac{1}{\sqrt{J}}$ we obtain $\frac{1}{J}\lambda_{\min}(\mathbf{I}_2^{(1)}(\delta_2)) \geq \lambda_{\min}(\widetilde{\mathbf{I}}_2(\delta_2))$ by a common refinement argument. Hence, $\lambda_{\min}(\mathbf{I}(\bar\delta_1 \otimes \delta_2)) = \lambda_{\min}(\widetilde{\mathbf{I}}_2(\delta_2))$ which proves (iii). □

For $q = 1$ we obtain that the symmetric design $\bar\delta_1 \otimes \delta_2^*$ is A-optimal if δ_2^* minimizes $\mathrm{trace}(\widetilde{\mathbf{I}}_2(\delta_2)) + J(J-1)\,\mathrm{trace}(\mathbf{I}_2^{(1)}(\delta_2))$. Note that the conditions on δ_2^* are weighted optimality criteria as introduced by Läuter (1974) in the situation of model uncertainty.

Example 1. An experimental situation in which different treatments $j = 1, ..., J$ can be applied at various dose levels x, $0 \leq x \leq b$, is described by the linear model $E(Y(j,x)) = \beta_0 + \beta_j x$ which fits into (1) with $f_0 = \mathbf{1}$ and $f_1(x) = x$.

By a majorization argument the optimal marginal design δ_2^* is concentrated on the endpoints of the interval $[0, b]$, i.e. $\delta_2^* \in \Delta_{0,b} = \{\delta_2;\ \delta_2(b) = w,\ \delta_2(0) = 1 - w\}$. It remains to determine the optimal weight $w^* = \delta_2^*(b)$ at the maximal dose b.

For $\delta_2 \in \Delta_{0,b}$ the eigenvalues of the information matrices in the relevant marginal models (4) and (5) are given by $\widetilde\lambda = \frac{1}{2}\left(1 + \frac{b^2}{J}w \pm \left((1 - \frac{b^2}{J}w)^2 + 4\frac{b^2}{J}w^2\right)^{1/2}\right)$ and $\lambda^{(1)} = b^2 w$, respectively. According to Theorem 1 the D-, A- and E-optimal weights w^* are calculated as $w_D^* = 1/(J+1)$, $w_A^* = J/(J+(J+b^2)^{1/2})$ and $w_E^* = 2J/(4J+b^2)$.

In an additive model without interactions, $f_1 = \mathbf{1}$, the conditions of Theorem 1 substantially simplify as $\mathbf{I}_2^{(1)}(\delta_2) = 1$ (see e.g. Schwabe, 1996, Corollary 6.15, for the A- and D-criterion)

Corollary 1. *In the additive model $E(Y(j,x)) = f_0(x)^\top\beta_0 + \beta_j$ the symmetric design $\bar\delta_1 \otimes \delta_2^*$ is Φ_q-optimal if δ_2^* is Φ_q-optimal in the weighted marginal model $E(Y_2(x)) = f_0(x)^\top\beta_0 + \frac{1}{\sqrt{J}}\beta_1$, for $0 \leq q \leq \infty$.*

As a consequence we recover the D-optimality of the product $\bar\delta_1 \otimes \delta_2^*$ of the D-optimal marginals in an additive model.

Example 2. One particular additive model is given by a one-way layout with additional regression $E(Y(j,x)) = \beta_j + \beta_0 x$, where $f_0(x) = x$. As in Example 1 the optimal marginal design δ_2^* is concentrated on the endpoints of the interval by a majorization argument, i. e. $\delta_2^* \in \Delta_{0,b}$ in case $\mathcal{X} = [0, b]$.

The eigenvalues are $\widetilde{\lambda} = \frac{1}{2J}\left(1 + Jb^2w \pm ((1 - Jb^2w)^2 + 4Jb^2w^2)^{1/2}\right)$ for $\delta_2 \in \Delta_{0,b}$ in the weighted marginal model. Consequently, the D-, A- and E-optimal weights are $w_D^* = 1/2$, $w_A^* = 1/((Jb^2+1)^{1/2}+1)$ and $w_E^* = 2/(Jb^2+4)$, respectively.

EXAMPLE 3. In a two-way layout $E(Y(i,j)) = \alpha_i + \beta_j$ two different kinds of treatments are simultaneously applied, $i = 1, ..., I$, $j = 1, ..., J$, which do not interact with each other. To identify the parameters we assume a control level I for the first factor, i. e. we impose the identifiability condition $\alpha_I = 0$. Here, the qualitative factor with levels $i = 1, ..., I$ plays the role of the second factor x in (1), and $f_0 = (\mathbf{1}_1, ..., \mathbf{1}_{I-1})^\top$ is a vector of indicator functions on $\{1, ..., I\}$.

Further symmetry considerations based on permutations of the actual treatment (non-control) levels $i = 1, ..., I-1$ shows that we can confine to the essentially complete class of marginal designs δ_2 on $\{1, ..., I\}$ with equal weights for those actual treatment levels, $\delta_2(1) = ... = \delta_2(I-1) = w$ and $\delta_2(I) = 1-(I-1)w$. For such designs $\widetilde{\lambda}_{1,2} = \frac{1}{2J}\left(1 + Jw \pm ((1-Jw)^2 + 4(I-1)Jw^2)^{1/2}\right)$ and $\widetilde{\lambda}_3 = ... = \widetilde{\lambda}_I = w$ are the eigenvalues in the weighted marginal model. The D-, A- and E-optimal weights w^* are $w_D^* = 1/I$, $w_A^* = 1/(I-1+(J+I-1)^{1/2})$ and $w_E^* = 2/(J+4(I-1))$.

For illustrative purposes we add another example with a more complicated interaction structure.

EXAMPLE 4. In the model $E(Y(j,x)) = \beta_{j1} + \beta_{j2}x + \beta_0 x^2$ the second factor shows a quadratic response, where the intercept and the slope depend on the treatments $j = 1, ..., J$ while the curvature is constant over the groups. Hence, the marginal regression functions are $f_0(x) = x^2$ and $f_1(x) = (1, x)^\top$.

(i) If the quantitative factor x varies over a symmetrical interval, $x \in [-b, b]$, with respect to some reference point $x_0 = 0$, then, by a symmetrization and majorization argument, the optimal marginal design δ_2^* is concentrated on $\{-b, 0, b\}$ with equal weights on the endpoints, $\delta_2^* \in \Delta_{-b,0,b}^{(\mathrm{symm})} = \{\delta_2;\ \delta_2(b) = \delta_2(-b) = w,\ \delta_2(0) = 1-2w\}$. $\widetilde{\lambda}_{1,2} = \frac{1}{2J}\left(1 + 2Jb^4w \pm ((1-2Jb^4w)^2 + 16Jb^4w^2)^{1/2}\right)$, $\widetilde{\lambda}_3 = \frac{2}{J}b^2w$ and $\lambda_1^{(1)} = 1$, $\lambda_2^{(1)} = 2b^2w$ are the eigenvalues for (4) and (5), respectively. Consequently, $w_D^* = \frac{1}{2}(J+1)/(J+2)$ and $w_A^* = \frac{1}{2}(J^2b^2+1)^{1/2}/((J^2b^2+1)^{1/2} + (Jb^4+1)^{1/2})$ are the D- and A-optimal weights. For the E-criterion the optimal weight is $w_E^* = 1/(Jb^4+4)$ if both the interval and the number of treatments are small, $Jb^2 \le 2$. For larger intervals or higher number of treatments, $Jb^2 \ge 2$, the smallest eigenvalue of $\widetilde{\mathbf{I}}_2(\delta_2^*)$ has multiplicity two, and the E-optimal weight is $w_E^* = \frac{1}{2}(Jb^2-1)/(Jb^4+Jb^2-b^2)$.

(ii) For general intervals, $x \in [a, b]$, we note that a translation in x preserves the structure of the model, up to a linear transformation of the regression functions.

Hence, the D-optimal design will be transformed accordingly, and the D-optimal weights are $\frac{1}{2}(J+1)/(J+2)$ at the endpoints and $(J-1)/(J+2)$ at the midpoint of that interval.

For various other interaction structures between a one-way layout and a quadratic response complete lists of D-, A- and E-optimal designs are presented in Schwabe (1996, p 94, and 1996b) in the special case of a standardized symmetric marginal design region $[-1, 1]$.

4. Symmetric subsystems of parameters.

In many situations the main interest is in parts of the parameters rather than in the whole parameter vector itself. This can be described by a linear functional ψ on β with $\psi(\beta) = L\beta$. Then the Φ_q-criteria aim at minimizing the "q-norm" of the eigenvalue $\nu_1, ..., \nu_{p(\psi)}$ of the covariance matrix $\mathbf{C}_\psi(\delta) = L\mathbf{I}(\delta)^- L^\top$ within the class of those design for which ψ is identifiable. In this case $\mathbf{I}(\delta)^-$ can be any arbitrary generalized inverse of $\mathbf{I}(\delta)$. We will treat symmetric parameter subsystems which are not affected by the permutations of the levels j of the qualitative factor besides a possible orthogonal transformation applied to ψ. These transformations preserve the eigenvalues of the covariance matrix $\mathbf{C}_\psi$ and the Φ_q-criteria are invariant with respect to permutations of j. Hence, we can confine to symmetric designs $\bar{\delta}_1 \otimes \delta_2$ and we will characterize the optimal marginal design δ_2^* by properties in the standard marginal model (3). As for the information matrices the covariance matrices in the marginal models are indicated by an additional subscript "$_2$".

First we consider the parameters $\bar{\beta} = (\beta_0^\top, \bar{\beta}_1^\top)^\top$ which are related to the mean response function $\frac{1}{J}\sum_{j=1}^J E(Y(j,x)) = f_0(x)^\top\beta_0 + f_1(x)^\top\bar{\beta}_1$ and, hence, $\bar{\beta}_1 = \frac{1}{J}\sum_{j=1}^J \beta_j$.

Theorem 2. $\mathbf{C}_{\bar{\beta}}(\delta) \geq \mathbf{C}_{\bar{\beta}}(\bar{\delta}_1 \otimes \delta_2)$ *for every design* δ, *where* δ_2 *is the second marginal of* δ.

REMARK. Here "$\geq$" denotes the usual uniform matrix ordering, i.e. $A \geq B$ if $A - B$ is positive semidefinite.

Proof of Theorem 2. For a symmetric design $\bar{\delta}_1 \otimes \delta_2$ a generalized inverse of the information matrix is given by

$$\mathbf{I}(\bar{\delta}_1 \otimes \delta_2)^- = \begin{pmatrix} C_0(\delta_2) & -\mathbf{1}_J^\top \otimes C_{01}(\delta_2) \\ -\mathbf{1}_J \otimes C_{01}^\top(\delta_2) & \mathbf{1}_J\mathbf{1}_J^\top \otimes C_1(\delta_2) + (J\mathbf{E}_J - \mathbf{1}_J\mathbf{1}_J^\top) \otimes \mathbf{I}^{(1)}(\delta_2)^- \end{pmatrix} \quad (8)$$

where C_0, C_{01}, and C_1 denote the blocks in an appropriately partitioned generalized inverse

$$\mathbf{I}_2(\delta_2)^- = \begin{pmatrix} C_0(\delta_2) & C_{01}(\delta_2) \\ C_{01}^\top(\delta_2) & C_1(\delta_2) \end{pmatrix}$$

of the information matrix in the standard marginal model (3) (see Schwabe, 1996, p 86, for the particular shape). In particular, if the components β_0 or β_1 are identifiable, then $C_0(\delta_2) = \mathbf{C}_{2,\beta_0}(\delta_2)$ and $C_1(\delta_2) = \mathbf{C}_{2,\beta_1}(\delta_2)$ are the corresponding covariance matrices.

From (8) we obtain that $\mathbf{C}_{\bar\beta}(\bar\delta_1 \otimes \delta_2) = \mathbf{C}_2(\delta_2)$. Now, the standard marginal model (3) can be regarded as a submodel of (1), at least, after some reparametrization which leaves $\bar\beta$ unchanged. Hence, by a common refinement argument $\mathbf{C}_{\bar\beta}(\delta) \geq \mathbf{C}_2(\delta_2)$, where δ_2 is the marginal of δ, which completes the proof. □

Note that the previous result can also be obtained by a straightforward argument that the parameter vector $\bar\beta$ associated with the mean response is invariant with respect to the permutations of j. In the sequel we will make extensive use of the representation (8) for the generalized inverse of the information matrix $\mathbf{I}(\bar\delta_1 \otimes \delta_2)$ obtained in the proof of Theorem 2. As a direct consequence we get the following characterization which relates the Φ_q-optimality for $\bar\beta$ and parts of it to properties in the standard marginal model (3).

Corollary 2. *The symmetric design $\bar\delta_1 \otimes \delta_2^*$ is Φ_q-optimal*
(i) *for $\bar\beta$ if δ_2^* is Φ_q-optimal.*
(ii) *for $\bar\beta_1$ if δ_2^* is Φ_q-optimal for β_1.*
(iii) *for $\bar\beta_0$ if δ_2^* is Φ_q-optimal for β_0.*

Next we are interested in the parameters associated with the interactions, $\psi(\beta) = (\beta_1^\top, ..., \beta_J^\top)^\top$, and we look for designs which are optimal for $\beta_1, ..., \beta_J$ (for ψ).

Lemma 2. *Let $\nu_1, ..., \nu_{p(1)}$ be the eigenvalues of the covariance matrix $\mathbf{C}_{2,\beta_1}(\delta_2)$ for β_1 in the standard marginal model, and let $\nu_1^{(1)}, ..., \nu_{p(1)}^{(1)}$ be the eigenvalues of $\mathbf{C}_2^{(1)}(\delta_2) = \mathbf{I}_2^{(1)}(\delta_2)^{-1}$. Then the eigenvalues of the covariance matrix $\mathbf{C}_{\beta_1,...,\beta_J}(\bar\delta_1 \otimes \delta_2)$ for $\psi(\beta) = (\beta_1^\top, ..., \beta_J^\top)^\top$ are given by $J\nu_1, ..., J\nu_{p(1)}$ and $J-1$ replicates of $J\nu_1^{(1)}, ..., J\nu_{p(1)}^{(1)}$.*

Proof. By (8) the covariance matrix

$$\mathbf{C}_{\beta_1,...,\beta_J}(\bar\delta_1 \otimes \delta_2) = \mathbf{1}_J \mathbf{1}_J^\top \otimes \mathbf{C}_{2,\beta_1},(\delta_2) + (J\mathbf{E}_J - \mathbf{1}_J\mathbf{1}_J^\top) \otimes \mathbf{C}_2^{(1)}(\delta_2)$$

is a combination of the associated covariance matrices in the marginal models. If z is an eigenvector of $\mathbf{C}_{2,\beta_1}(\delta_2)$ associated with the eigenvalue ν_i, then $\mathbf{1}_J \otimes z$ is an eigenvector of $\mathbf{C}_{\beta_1,...,\beta_J}(\bar\delta_1 \otimes \delta_2)$ with eigenvalue $J\nu_i$. For the remaining eigenvalue we conclude as in the proof of Lemma 1 that $\ell^\top \otimes z_{(1)}$ is an eigenvector of $\mathbf{C}_{\beta_1,...,\beta_J}(\bar\delta_1 \otimes \delta_2)$

with eigenvalue $J\nu_i^{(1)}$ if ℓ is a contrast and $z_{(1)}$ is an eigenvector of $\mathbf{C}_2^{(1)}(\delta_2)$ associated with the eigenvalue $\nu_i^{(1)}$. □

Denote by and Φ_{2,q,β_1} the Φ_q-criterion function for β_1 in the standard marginal model (3).

Theorem 3. *The symmetric design* $\bar{\delta}_1 \otimes \delta_2^*$ *is*

(i) Φ_q*-optimal for* $\beta_1, ..., \beta_J$ *if* δ_2^* *minimizes* $\Phi_{2,q,\beta_1}(\delta_2) + (J-1)\Phi_{2,q}^{(1)}(\delta_2)$*, for* $0 < q < \infty$.

(ii) *D-optimal for* $\beta_1, ..., \beta_J$ *if* δ_2^* *minimizes* $\det(\mathbf{C}_{2,\beta_1}(\delta_2)) \det(\mathbf{C}_2^{(1)}(\delta_2))^{J-1}$.

(iii) *E-optimal for* $\beta_1, ..., \beta_J$ *if* δ_2^* *is E-optimal for* β_1 *in the standard marginal model (3).*

Proof. (i) and (ii) follow directly from Lemma 2. For (iii) we have to note, again, that $\lambda_{\max}(\mathbf{C}_2^{(1)}(\delta_2)) \leq \lambda_{\max}(\mathbf{C}_{2,\beta_1}(\delta_2))$ by a refinement argument. □

In the case $q = 1$ we obtain that the symmetric design $\bar{\delta}_1 \otimes \delta_2^*$ is A-optimal for $\beta_1, ..., \beta_J$ if δ_2^* minimizes $\mathrm{trace}(\mathbf{C}_{2,\beta_1}(\delta_2)) + (J-1)\,\mathrm{trace}(\mathbf{C}_2^{(1)}(\delta_2))$.

Example 5. In the setting of the treatment dose model of Example 1 the main interest is in the treatment effects $\beta_1, ..., \beta_J$ which represent the slopes corresponding to the dose effects of each single treatment. Again, we can confine to the essentially complete class $\Delta_{0,b}$ of marginal designs concentrated on a zero and a maximal dose level. The marginal covariance matrices $\mathbf{C}_{2,\beta_1}(\delta_2) = (b^2 w(1-w))^{-1}$ and $\mathbf{C}_2^{(1)}(\delta_2) = (b^2 w)^{-1}$ reduce to one-dimensional quantities, from which the optimal weights $w_D^* = J/(J+1)$, $w_A^* = \sqrt{J}/(\sqrt{J}+1)$ and $w_E^* = \frac{1}{2}$ are obtained which produce a D-, A- and E-optimal design, respectively.

Example 6. In the setting of Example 4 the intercept and the slope of a quadratic covariate are influenced by the treatment levels. The associated parameters are of particular interest, here.

(i) On a symmetric design region $\mathcal{X} = [-b, b]$ we can confine, again, to the class $\Delta_{-b,0,b}^{(\mathrm{symm})}$ of symmetric three-point designs. The eigenvalues of the marginal covariance matrices $\mathbf{C}_{2,\beta_1}$ and $\mathbf{C}_2^{(1)}$ associated with the treatment effects are found to be $\nu_1 = (1-2w)^{-1}$, $\nu_2 = \nu_2^{(1)} = (2b^2 w)^{-1}$ and $\nu_1^{(1)} = 1$. This gives D- and A-optimal weights $w_D^* = \frac{1}{2}J/(J+1)$ and $w_A^* = \frac{1}{2}\sqrt{J}/(\sqrt{J}+b)$ for the treatment effects $\beta_1, ..., \beta_J$. For the E-criterion the two cases of small and large intervals have to be distinguished: $w_E^* = (b^4+4)^{-1}$ for $b^2 \leq 2$ and $w_E^* = \frac{1}{2}b^{-4}(b^2-1)$ for $b^2 \geq 2$.

(ii) For the asymmetric design region $[0, b]$ we notice that in contrast to the inference on the full parameter vector the D-optimal designs cannot be obtained by translations from the symmetric interval. By means of the Kiefer-Wolfowitz equivalence theorem the optimal marginal design δ_2^* can be found in the class $\Delta_{0,x,b}$

J	1	2	3	4	5	10	20	50	...	∞
w_0^*	0.500	0.470	0.468	0.470	0.472	0.482	0.490	0.495	...	0.500
w_b^*	0.073	0.219	0.289	0.334	0.364	0.428	0.463	0.485	...	0.500
w_x^*	0.427	0.311	0.243	0.196	0.164	0.090	0.047	0.020	...	0.000
x^*/b	0.414	0.486	0.506	0.510	0.510	0.508	0.505	0.502		

Table 1: Weights and locations of the interior design points of D-optimal designs in a one-way layout with quadratic covariate on the interval $[0, b]$: inference on the interaction parameters

of three-point designs with $\delta_2(0) = w_0$, $\delta_2(b) = w_b$ and $\delta_2(x) = w_x = 1 - w_0 - w_b$ for some x, $0 < x < b$. Numerical solutions to this problem are given in Table 1.

Finally, in a model without interactions the optimality for β_1 in the marginal directly carries over to the optimality for the treatment effects $\beta_1, ..., \beta_J$

Corollary 3. *In the additive model $E(Y(j,x)) = f_0(x)^T\beta_0 + \beta_j$ the symmetric design $\bar{\delta}_1 \otimes \delta_2^*$ is Φ_q-optimal for $\beta_1, ..., \beta_J$ simultaneously in $q \in [0, \infty]$, if δ_2^* is optimal for the constant term β_1 in the marginal model $E(Y_2(x)) = f_0(x)^\top\beta_0 + \beta_1$.*

5. Explicit mean function.

In contrast to the previous sections we consider now an extended form

$$E(Y(j,x)) = f(x)^\top\bar{\beta} + f_1(x)^\top\beta_j \tag{9}$$

of the model (1) in which an explicit mean function $f(x)^\top\bar{\beta} = \frac{1}{J}\sum_{j=1}^J E(Y(j,x))$ is present, and where $f = (f_0^\top, f_1^\top)^\top$ and $\bar{\beta} = (\beta_0^\top, \bar{\beta}_1^\top)^\top$ are defined as before. The identifiability of $\beta = (\bar{\beta}^\top, \beta_1^\top, ..., \beta_J^\top)^\top$ is accomplished by the natural side condition $\sum_{j=1}^J \beta_j = \mathbf{0}$, which makes $f_1(x)^\top\beta_j$ to the treatment effect of level j compared to the mean response $f(x)^\top\bar{\beta}$.

As the indentification condition is invariant with respect to the permutations of the levels j the Φ_q-criteria remain to be invariant also in the model (9), $0 < q \leq \infty$. It is straightforward from (8) that

$$\mathbf{C}(\bar{\delta}_1 \otimes \delta_2^*) = \begin{pmatrix} \mathbf{C}_2(\delta_2) & \mathbf{0} \\ \mathbf{0} & (J\mathbf{E}_J - \mathbf{1}_J\mathbf{1}_J^\top) \otimes \mathbf{C}_2^{(1)}(\delta_2) \end{pmatrix} \tag{10}$$

is the covariance matrix of a symmetric design in the present extended model.

Theorem 4. *The symmetric design* $\bar\delta_1 \otimes \delta_2^*$ *is*
(i) Φ_q*-optimal in (9) if* δ_2^* *minimizes* $\Phi_{2,q}(\delta_2) + (J-1)J^q\Phi_{2,q}^{(1)}(\delta_2)$, $0 < q < \infty$.
(ii) E*-optimal in (9) if* δ_2^* *maximizes* $\min(\lambda_{\min}(\mathbf{I}_2(\delta_2)), \frac{1}{J}\lambda_{\min}(\mathbf{I}_2^{(1)}(\delta_2)))$.

Proof. The non-zero eigenvalues of the covariance matrix $\mathbf{C}_{\beta_1,\ldots,\beta_J}(\bar\delta_1 \otimes \delta_2) = (J\mathbf{E}_J - \mathbf{1}_J\mathbf{1}_J^\top) \otimes \mathbf{C}_2^{(1)}(\delta_2)$ are $J-1$ replicates of $J\nu_1^{(1)}, \ldots, J\nu_{p(1)}^{(1)}$ as can be seen from the proof of Lemma 2. Hence, by the block diagonal structure of (10) the result is immediate. □

For $q = 1$ we obtain that the symmetric design $\bar\delta_1 \otimes \delta_2^*$ is A-optimal if δ_2^* minimizes $\mathrm{trace}(\mathbf{I}_2(\delta_2)^{-1}) + J(J-1)\,\mathrm{trace}(\mathbf{I}_2^{(1)}(\delta_2)^{-1})$.

Example 7. The model $E(Y(j,x)) = \beta_{01} + \beta_{02}x + \beta_j x$ is the extended form of the treatment dose model of Example 1. Again, $\delta_2^* \in \Delta_{0,b}$, and the A-optimal weight is $w_A^* = (J^2-J+1)^{1/2}/((J^2-J+1)^{1/2} + (b^2+1)^{1/2})$. For the E-criterion the optimal weight is given by $w_E^* = J(J-1)/(J^2 + Jb^2 - b^2)$, for $J \geq 2$. Note that for $J = 2$ the optimal weight $w_E^* = 2(b^2+4)^{-1}$ is the same as in the marginal model.

Example 8. For the quadratic response model of Example 4 with an explicit mean response the A- and E-optimal weights of $\delta_2^* \in \Delta_{-b,0,b}^{(\mathrm{symm})}$ are given by $w_A^* = \frac{1}{2}((J^2-J+1)b^2+1)^{1/2}/\left(((J^2-J+1)b^2+1)^{1/2} + (b^4+1)^{1/2}\right)$, $w_E^* = (b^4+4)^{-1}$ if $Jb^2 \leq 2$, and $w_E^* = \frac{1}{2}J(Jb^2-1)/(J^2b^2 + Jb^4 - b^2)$ if $Jb^2 \geq 2$.

In case of no interactions the optimization problem can be completely reduced to the marginal model as follows

Corollary 4. *In the additive model* $E(Y(j,x)) = f(x)^\top\bar\beta_0 + \beta_j$ *with explicit mean,* $f = (f_0^\top, \mathbf{1})^\top$*, the symmetric design* $\bar\delta_1 \otimes \delta_2^*$ *is* Φ_q*-optimal if* δ_2^* *is* Φ_q*-optimal in the marginal model* $E(Y_2(x)) = f_0(x)^\top\beta_0 + \beta_1$*, for* $0 < q \leq \infty$.

Remark. For the treatment effects compared to the mean response it follows directly from (10) that the symmetric design $\bar\delta_1 \otimes \delta_2^*$ is Φ_q-optimal for $\beta_1, \ldots, \beta_J$ if δ_2^* is Φ_q-optimal in the restricted marginal model (5).

6. Higher dimensions.

The previous results can be extended to K-way layouts

$$E(Y(j_1,\ldots,j_K,x)) = f(x)^\top\bar\beta + \sum_{k=1}^K f_k(x)^\top\beta_{j_k}^{(k)}, \tag{11}$$

$j_k = 1,\ldots,J_k$, $k = 1,\ldots,K$, with covariates, $x \in \mathcal{X}$, and with an explicit mean function $f(x)^\top\bar\beta = \frac{1}{J_K\cdot\ldots\cdot J_K}\sum_{j_1=1}^{J_1}\ldots\sum_{j_K=1}^{J_K} E(Y(j_1,\ldots,j_K,x))$. The f_k are the regression functions associated with the kth restricted marginal model $E(Y_2^{(k)}(x)) = f_k(x)^\top\beta_k$

in the factor x, and the components of f_k are assumed to be contained in the space spanned by the components of f. Furthermore the natural side condition $\sum_{j=1}^{J_k} \beta_j^{(k)} = \mathbf{0}$ is imposed on each component to indentify the parameters.

Again, symmetry considerations show that each design δ is dominated by its symmetrization $\bar{\delta}_{1,...,K} \otimes \delta_2$ with respect to every Φ_q-criterion, where $\bar{\delta}_{1,...,K}$ is the uniform design on all level combinations $(j_1, ..., j_K)$ of the K-way layout and δ_2 is the marginal of δ for the factor x. The covariance matrix $\mathbf{C}(\bar{\delta}_{1,...,K} \otimes \delta_2)$ is block diagonal, and Theorem 4 can repeatedly applied to each of the K qualitative factors. For example, if δ_2^* minimizes $\mathrm{trace}(\mathbf{I}_2(\delta_2)^{-1}) + \sum_{k=1}^{K} J_k(J_k - 1)\,\mathrm{trace}(\mathbf{I}_2^{(k)}(\delta_2)^{-1})$, where $\mathbf{I}_2^{(k)}(\delta_2) = \int f_k f_k^\top d\delta_2$ is the information matrix in the kth restricted marginal model, then the symmetric design $\bar{\delta}_{1,...,K} \otimes \delta_2^*$ is A-optimal.

For the D-criterion a minimal reparametrization of the model (11) has to be considered or, equivalently, the product of the non-zero eigenvalues of the covariance matrix $\mathbf{C}(\bar{\delta}_{1,...,K} \otimes \delta_2)$ is to be minimized. With this approach a D-optimal symmetric design $\bar{\delta}_{1,...,K} \otimes \delta_2^*$ is obtained if δ_2^* maximizes $\det(\mathbf{I}_2(\delta_2)) \prod_{k=1}^{K} \det(\mathbf{I}_2^{(k)}(\delta_2))^{J_k - 1}$.

Again, for the additive model $E(Y(j_1, ..., j_k, x)) = f_0(x)^\top \beta_0 + \mu_0 + \sum_{k=1}^{K} \beta_{j_k}^{(k)}$ the situation substantially simplifies, and the symmetric design is Φ_q-optimal if its marginal associated with the factor x is Φ_q-optimal in the corresponding marginal model $E(Y_2(x)) = f_0(x)^\top \beta_0 + \mu_0$.

References

Atkinson, A. C. and A. N. Donev (1992). *Optimum Experimental Designs.* Clarendon Press, Oxford.

Bandemer, H. (ed.) (1977). *Theorie and Anwendung der optimalen Versuchsplanung I. Handbuch zur Theorie.* Akademie-Verlag, Berlin.

Kiefer, J. (1974). General equivalence theory for optimum designs (approximate theory). *Ann. Statist.* **2**, 849–879.

Kurotschka, V., R. Schwabe and W. Wierich (1992). Optimum designs for partly interacting qualitative and quantitative factors. In A. Pázman and J. Volaufová (eds.), *Probastat '91, Proc. Int. Conf. Probab. Math. Statist., Bratislava 1991.* Slovak Academy of Sciences, Bratislava, 102–108.

Läuter, E. (1974). Experimental design in a class of models. *Math. Operationsforsch. Statist.* **5**, 379–398.

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

Schwabe, R. (1996). *Optimum Designs for Multi-factor Models. Lecture Notes in Statistics* **113**. Springer, New York.

Schwabe, R. (1996b). *E*-optimum designs in linear models with both qualitative and quantitative factors of influence. In S. M. Ermakov and V. B. Melas (eds.), *Mathematical Methods in Stochastic Simulation and Experimental Design. Proc. 2nd St. Petersburg Workshop on Simulation, 1996.* Publishing House of Saint Petersburg University, Saint Petersburg, 256–260.