OPTIMAL DESIGNS FOR COMPLETE INTERACTION STRUCTURES

By Rainer Schwabe

Freie Universität Berlin

In linear models with complete interaction structures product type designs are shown to be Φ_q-optimal for the maximal interactions, $0 \le q \le \infty$, and D-optimal simultaneously for the whole parameter vector and for testing against any complete submodel.

1. Introduction. We consider a general multi-factor linear model

$$E(X(t_1, \dots, t_K)) = \mu_{\mathcal{H}}(t_1, \dots, t_K) = \textstyle\sum_{H \in \mathcal{H}} \bigotimes_{k \in H} f_k(t_k)^T \beta_H,$$

$t_k \in T_k$, $k = 1, \dots, K$, with interaction structure $\mathcal{H} \subseteq \mathcal{P}(\{1, \dots, K\}) := \{H; H \subseteq \{1, \dots, K\}\}$. Throughout this paper we assume that the interaction structure $\mathcal{H}$ is complete, i. e. each interaction $H \in \mathcal{H}$ comes with all its lower order interactions:

Definition 1.1. $\mathcal{H}$ is complete if $\mathcal{P}(H) \subseteq \mathcal{H}$ for every $H \in \mathcal{H}$.

Models of these type include product models, $\mathcal{H} = \mathcal{P}(\{1, \dots, K\})$, considered by Hoel (1965), additive models, $\mathcal{H} = \{\emptyset, \{1\}, \dots, \{K\}\}$, (see e. g. Schwabe and Wierich, 1995) and complete M-factor interaction models, $\mathcal{H} = \{H; H \subseteq \{1, \dots, K\}, |H| \le M\}$, (see e. g. Schwabe, 1995). Particular examples will be given in section 4.

The aim of the present paper is to establish that, in the setting of approximate designs (see e. g. Kiefer, 1974), product type designs are simultaneously optimal for higher order interactions if their components are optimal in the corresponding marginal models $\mu_k(t_k) = \beta_0 + f_k(t_k)^T \beta_k$, $t_k \in T_k$.

Definition 1.2. $H \in \mathcal{H}$ is maximal in $\mathcal{H}$ if $H' \notin \mathcal{H}$ for every $H' \supset H$, $H' \ne H$.

We will show that the product type design $\delta^* = \bigotimes_{k=1}^K \delta_k^*$ is Φ_q-optimal for maximal interactions β_H, i. e. δ^* minimizes the q-"norm" of the eigenvalues of the covariance matrix of the least squares estimator $\hat{\beta}_H$ of β_H, if its components δ_k^* are Φ_q-optimal for β_k in the marginal models, $0 \le q \le \infty$. Note that the Φ_q-criteria include the common A-optimality, $q = 1$, and the D- and E-optimality as limiting cases, $q \to 0$ and $q \to \infty$, respectively. Moreover, for the D-criterion, the product type

The research was supported by grant Ku719/2-1 of the Deutsche Forschungsgemeinschaft.

AMS 1991 *subject classification.* 62K05.

Key words and phrases. Product type designs, multi-factor linear models, Φ_q-optimality, D-optimality, complete interactions, maximal interaction.

Abbreviated title. Designs for interactions.

design δ^* will be shown to be simultaneously optimal for the whole parameter vector $\beta = (\beta_H)_{H\in\mathcal{H}}$ as well as for testing every complete model $\mu_{\mathcal{H}}$ against every complete submodel $\mu_{\mathcal{H}^*}$, $\mathcal{H}^* \subset \mathcal{H}$.

2. Notations and auxiliary results. We recall that a subset $\beta_H = L_H\beta$ of parameters is identifiable under a design δ in a linear model $\mu(t) = a(t)^T\beta$ if the rows of the selection matrix L_H are contained in the space spanned by the regression function a evaluated at the support points t of δ or, equivalently, spanned by the columns of the information matrix $\mathbf{I}(\delta) = \int aa^T\, d\delta$. In the present setting the regression functions considered are $a_{\mathcal{H}}(t_1, ..., t_K) = (\bigotimes_{k\in H} f_k(t_k))_{H\in\mathcal{H}}$.

If β_H is identifiable the corresponding covariance matrix is given by $\mathbf{C}(\beta_H;\delta) = L_H\mathbf{I}(\delta)^- L_H^T$. For larger subsets $\beta_{\mathcal{H}'} = (\beta_H)_{H\in\mathcal{H}'}$ we have $L_{\mathcal{H}'} = \mathrm{diag}(L_H)_{H\in\mathcal{H}'}$ and $\mathbf{C}(\beta_{\mathcal{H}'};\delta) = L_{\mathcal{H}'}\mathbf{I}(\delta)^- L_{\mathcal{H}'}^T$ if each β_H, $H \in \mathcal{H}'$, is identifiable under δ. In case $\mathcal{H}' = \mathcal{H}$ we write $\mathbf{C}(\delta) = \mathbf{C}(\beta;\delta)$, for short.

In the corresponding marginal models the parameters are $\beta^{(k)} = (\beta_0, \beta_k^T)^T$ and the information matrices $\mathbf{I}_k(\delta_k) = \int a_k a_k^T\, d\delta_k$ where $a_k = (1, f_k^T)^T$. The marginal covariance matrices for $\beta^{(k)}$ and $\beta_k = L_k\beta^{(k)}$ are denoted by $\mathbf{C}_k(\delta_k) = \mathbf{I}_k(\delta_k)^{-1}$ and $\mathbf{C}_k(\beta_k;\delta_k) = L_k\mathbf{C}(\delta_k)L_k^T$ respectively. Note that β_k is identifiable under δ_k iff $\mathbf{I}_k(\delta_k)$ is regular.

As usual $\lambda_{\max}$ and tr denote the largest eigenvalue and the trace of a matrix.

For every pair of designs δ_H on $\times_{k\in H}T_k$ and $\delta_{\bar{H}}$ on $\times_{k\notin H}T_k$ the resulting product design on $\times_{k=1}^K T_k$ is denoted by $\pi_H(\delta_H, \delta_{\bar{H}})$.

Lemma 2.1. *Let* $\mu(t) = f_0(t)^T\beta_0 + f_1(t)^T\beta_1 + f_2(t)^T\beta_2$ *and let* $\mu_0(t) = f_0(t)^T\beta_0 + f_1(t)^T\beta_1$ *be a submodel of* μ.

(i) *If* β_1 *is identifiable under* δ *in* μ, *then* $\mathbf{C}_0(\beta_1;\delta) \le \mathbf{C}(\beta_1;\delta)$.

(ii) *If* β_1 *and* β_2 *are identifiable under* δ *in* μ, *then* $\det \mathbf{C}(\beta_1,\beta_2;\delta) = \det \mathbf{C}_0(\beta_1;\delta)\ \det \mathbf{C}(\beta_2;\delta)$.

Proof. (i) is a common refinement statement (see e. g. Kunert, 1983).

For (ii) we partition the information matrix

$$\mathbf{I}(\delta) = \begin{pmatrix} \mathbf{I}_0(\delta) & I_{12} \\ I_{12}^T & I_2 \end{pmatrix}$$

in μ according to the submodel, and the covariance matrix

$$\mathbf{C}(\beta_1,\beta_2;\delta) = \begin{pmatrix} C_1 & C_{12} \\ C_{12}^T & C_2 \end{pmatrix}$$

according to the components β_1 and β_2, $C_j = \mathbf{C}(\beta_j;\delta)$. Further let L be the selection matrix for $\beta_1 = L(\beta_0^T, \beta_1^T)^T$ in μ_0.

By the formula for generalized inverses of partitioned matrices

$$\begin{aligned} C_1 &= L(\mathbf{I}_0(\delta)^- + \mathbf{I}_0(\delta)^- I_{12}C_2I_{12}^T\mathbf{I}_0(\delta)^-)L^T \\ C_{12} &= L\mathbf{I}_0(\delta)^- I_{12}C_2 \end{aligned}$$

and hence

$$C_1 - C_{12}C_2^{-1}C_{12}^T = L\mathbf{I}_0(\delta)^- L^T = \mathbf{C}_0(\beta_1;\delta).$$

The result now follows from $\det \mathbf{C}(\beta_1,\beta_2;\delta) = \det(C_1 - C_{12}C_2^{-1}C_{12}^T)\ \det C_2$. □

REMARK 2.1. If $f_0 \equiv 1$ (or $f_0 \equiv 0$), then $\det \mathbf{C}(\delta) = \det \mathbf{C}_0(\delta)\ \det \mathbf{C}(\beta_2;\delta)$ from Lemma 2.1.

3. Optimality of product type designs. The first result deals with the optimality of product type designs for the highest interaction in total interactions models

THEOREM 3.1. *Let $\mathcal{H} = \mathcal{P}$ and let $H = \{1,\ldots,K\}$. If δ_k^* is Φ_q-optimal for β_k in the kth marginal model, $k = 1,\ldots,K$, then $\delta^* = \bigotimes_{k=1}^K \delta_k^*$ is Φ_q-optimal for β_H.*

PROOF. Note that $\beta_H = L\beta$ where $L = \bigotimes_{k=1}^K L_k$ and $\beta_k = L_k\beta^{(k)}$. Hence the covariance matrix $\mathbf{C}(\beta_H;\delta^*) = \bigotimes_{k=1}^K \mathbf{C}_k(\beta_k;\delta_k^*)$ factorizes. The Φ_q-optimality can now be obtained by means of the appropriate equivalence theorems. For $q < \infty$ see Rafajłowicz and Myszka (1988). For the E-citerion ($q = \infty$) a more implicit equivalence theorem (Pukelsheim, 1993, p. 182) has to be used: δ_k^* is E-optimal for β_k iff there exists a non-negative definite, symmetric matrix M_k, $\operatorname{tr} M_k = 1$, such that

$$a_k(t_k)^T\mathbf{C}_k(\delta_k^*)L_k^T\mathbf{C}_k(\beta_k;\delta_k^*)^{-1}M_k\mathbf{C}_k(\beta_k;\delta_k^*)^{-1}L_k\mathbf{C}_k(\delta_k^*)a_k(t_k) \le 1/\lambda_{\max}(\mathbf{C}_k(\beta;\delta_k^*))$$

for all $t_k \in T_k$.

Let $M = \bigotimes_{k=1}^K M_k$. Then M is non-negative definite, symmetric, $\operatorname{tr} M = 1$, and

$$a(t)^T\mathbf{C}(\delta^*)L^T\mathbf{C}(\beta_H;\delta^*)^{-1}M\mathbf{C}(\beta_H;\delta^*)^{-1}L\mathbf{C}(\delta^*)a(t) \le 1/\lambda_{\max}(\mathbf{C}(\beta_H;\delta^*)).$$

Thus the equivalence theorem proves the E-optimality of δ^* for β_H. □

Note that $H = \{1,\ldots,K\}$ is the only maximal element in $\mathcal{H} = \mathcal{P}$. For general, complete interaction structures $\mathcal{H}$ a maximal interaction H will be a proper subset of $\{1,\ldots,K\}$. Then, conditions are only required on those factors k belonging to H.

THEOREM 3.2. *Let $\mathcal{H}$ be complete, and let H be maximal in $\mathcal{H}$. If δ_k^* is Φ_q-optimal for β_k in the kth marginal model, $k \in H$, then $\delta^* = \pi_H(\bigotimes_{k\in H}\delta_k^*, \delta_{\bar H})$ is Φ_q-optimal for β_H, for every design $\delta_{\bar H}$ on $\times_{k\notin H}T_k$.*

PROOF. Consider the model $\mu_{\mathcal{P}(H)}$ of total interactions in the factors involved in H. Then $\mu_{\mathcal{P}(H)}$ can be regarded as a submodel of $\mu_{\mathcal{H}}$. Moreover, as H is maximal,

$$\mathbf{C}_{\mathcal{P}(\mathcal{H})}(\beta_H;\bigotimes_{k\in H}\delta_k^*) = \mathbf{C}(\beta_H;\delta^*),$$

and by Theorem 3.1 the design $\bigotimes_{k\in H}\delta_k^*$ is Φ_q-optimal for β_H in $\mu_{\mathcal{P}(H)}$. Hence, the result follows by a refinement argument (cf Lemma 2.1 (i)). □

The main result on Φ_q-optimality is the following extension to more than one maximal interaction:

THEOREM 3.3. *Let $\mathcal{H}$ be complete, $\mathcal{H}' \subseteq \{H; H$ maximal in $\mathcal{H}\}$. If δ_k^* is Φ_q-optimal for β_k in the kth marginal model, $k \in H(\mathcal{H}') := \bigcup_{H\in\mathcal{H}'} H$, then $\delta^* = \pi_{H(\mathcal{H}')}(\bigotimes_{k\in H(\mathcal{H}')}\delta_k^*, \delta_{\bar{H}(\mathcal{H}')})$ is Φ_q-optimal for $\beta_{\mathcal{H}'}$, for every design $\delta_{\bar{H}(\mathcal{H}')}$ on $\times_{k\notin H(\mathcal{H}')}T_k$.*

PROOF. The key observation is that the covariance matrix is block diagonal:

$$\mathbf{C}(\beta_{\mathcal{H}'};\delta^*) = \mathrm{diag}(\mathbf{C}(\beta_H;\delta^*))_{H\in\mathcal{H}'}$$

because each $H \in \mathcal{H}'$ is maximal. By Fan's (1954) result on the eigenvalues of block matrices the Φ_q-optimality of δ^* can be deduced for $\beta_{\mathcal{H}'}$ from the Φ_q-optimality of δ_k^* for every β_H, $H \in \mathcal{H}'$, established in Theorem 3.2. □

For additive models, $\mathcal{H} = \{\emptyset, \{1\}, \ldots, \{K\}\}$ all direct effects $\{k\}$ are maximal, and we revover a special case of Theorem 3.2 by Schwabe (1994). As indicated there no general results can be expected for Φ_q-optimality, $q > 0$, of product type designs if larger parts of the parameter vector are of interest including interactions which are not maximal. However, for D-optimality ($q = 0$) a stronger result can be obtained.

THEOREM 3.4. *Let $\mathcal{H}$ be complete, and let $\mathcal{H}' \subseteq \mathcal{H}$, $\mathcal{H}\setminus\mathcal{H}'$ complete. If δ_k^* is D-optimal in the kth marginal model, $k \in H(\mathcal{H}')$, then $\delta^* = \pi_{H(\mathcal{H}')}(\bigotimes_{k\in H(\mathcal{H}')}\delta_k^*, \delta_{\bar{H}(\mathcal{H}')})$ is D-optimal for $\beta_{\mathcal{H}'}$, for every design $\delta_{\bar{H}(\mathcal{H}')}$ on $\times_{k\notin H(\mathcal{H}')}T_k$.*

PROOF. First note that δ_k^* is D-optimal iff δ_k^* is D-optimal for β_k (see e.g. Schwabe and Wierich, 1995, Lemma 6). Hence δ^* is D-optimal for any set of maximal interactions by Theorem 3.3.

The proof will be completed by a finite induction: Therefore assume that the theorem is valid for $\mathcal{H}''$. Let H be maximal in $\mathcal{H} \setminus \mathcal{H}''$. Then it suffices to show that the theorem is also valid for $\mathcal{H}'' \cup \{H\}$, because every set $\mathcal{H}'$ can be generated successively by adding relatively maximal interactions. In fact, by Lemma 2.1 (ii),

$$\begin{aligned}\det \mathbf{C}(\beta_{\mathcal{H}''\cup\{H\}};\delta) &= \det \mathbf{C}(\beta_{\mathcal{H}''};\delta)\ \det \mathbf{C}_{\mathcal{H}\setminus\mathcal{H}''}(\beta_H;\delta)\\ &\leq \det \mathbf{C}(\beta_{\mathcal{H}''};\delta^*)\ \det \mathbf{C}_{\mathcal{H}\setminus\mathcal{H}''}(\beta_H;\delta^*),\end{aligned}$$

where the inequality follows by Theorem 3.2 and the assumption. □

In particular, for $\mathcal{H}' = \mathcal{H}$, the design $\bigotimes_{k=1}^K \delta_k^*$ is D-optimal if each of its components δ_k^* is D-optimal in the corresponding marginal model. Finally, note that the full product type design $\bigotimes_{k=1}^{K} \delta_k^*$ is D-optimal for $\beta_{\mathcal{H}'}$ simultaneously in $\mathcal{H}$ and $\mathcal{H}' \subseteq \mathcal{H}$ when $\mathcal{H}$ and $\mathcal{H} \setminus \mathcal{H}'$ are complete.

4. Examples.

4.1. Multilinear regression on $[-1,1]^K$. The full 2^K factorial design is simultaneously Φ_q-optimal for every set of maximal interactions in every complete interactions model

$$\mu_{\mathcal{H}}(t_1, \ldots, t_K) = \textstyle\sum_{H \in \mathcal{H}} \beta_H \prod_{k \in H} t_k.$$

Moreover, the 2^K factorial is D-optimal for the whole parameter vector as well as for any set $\mathcal{H}'$ of higher interactions, such that the remaining interaction structure $\mathcal{H} \setminus \mathcal{H}'$ is complete. In particular, in the model of complete M-factor interactions the 2^K factorial is the unique D-optimal design (the unique D-optimal design for all interactions of m up to M factors, i. e. $\mathcal{H}' = \{H;\ H \subseteq \{1, \ldots, K\},\ m \leq |H| \leq M\}$; the unique Φ_q-optimal design for all M-factor interactions) if $2M \geq K$. In case $2M < K$ also 2^d fractions may be optimal, $2M \leq d < K$. For example, in the additive model ($M = 1$) there exist optimal 2^d fractional factorials for $d \leq \log_2(K + 1)$, $K \geq 3$, constructed from Hadamard matrices.

4.2. K-way layout. Every complete K-way layout

$$\mu_{\mathcal{H}}(i(1), \ldots, i(K)) = \textstyle\sum_{H \in \mathcal{H}} \beta^{(H)}_{H(i(1),\ldots,i(K))},$$

$i(k) = 1, \ldots, I_k$, $k = 1, \ldots, K$, where $H(i(1), \ldots, i(K)) = (i(k_1), \ldots, i(k_{|H|}))$ is the projection of $(i(1), \ldots, i(K))$ onto the components $k_1, \ldots, k_{|H|}$ of H, can be reparametrized to the form considered above, if suitable identifiability conditions $\sum \beta_{i(k_1),\ldots,i(k_{|H|})} = 0$, $H \in \mathcal{H}$, are imposed. (Here the summation is to be considered over every set of $|H| - 1$ components). Hence the equireplicated design is D-optimal for testing every model with complete interaction structure $\mathcal{H}$ against each submodel with complete interactions structure $\mathcal{H} \setminus \mathcal{H}'$.

4.3. Qualitative and quantitative factors. Following Kurotschka (1988) those models are of particular interest in applications which contain both kinds of factors: qualitative ones which may vary only over a finite number of levels (see e. g. 4.2) and quantitative ones which vary over a whole continuum of levels (see e. g. 4.1). To be more specific, in a two-way layout with additional straight line regression on $[-1, 1]$ the model may be either additive

$$\mu(i, j, t) = \alpha_i^{(1)} + \alpha_j^{(2)} + \beta t$$

or may have total interactions

$$\mu(i, j, t) = \alpha_{ij} + \beta_{ij} t$$

or, as indermediate models, the additive model is supplemented by some or all of the two-factor interactions α_{ij}, $\beta_i^{(1)} t$ and $\beta_j^{(2)} t$. For all those 9 different models the equireplicated design on $\{1,\ldots,I\}\times\{1,\ldots J\}\times\{-1,1\}$ is D-optimal for the whole parameter vector (under suitable identifiability conditions) as well as for estimating (testing) the higher interaction terms.

REFERENCES

FAN, K. (1954). Inequalities for eigenvalues of Hermitian matrices. *Nat. Bur. Standards, Appl. Math. Ser.* **39** 131–139.

HOEL, P. G. (1965). Minimax designs in two dimensional regression. *Ann. Math. Statist.* **36** 1097–1106.

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

KUNERT, J. (1983). Optimal design and refinement of the linear model with applications to repeated measurement designs. *Ann. Statist.* **11** 247–257.

KUROTSCHKA, V. G. (1988). Characterizations and examples of optimal experiments with qualitative and quantitative factors. In *Model-Oriented Data Analysis, Proceedings Eisenach 1987.* (V. Fedorov and H. Läuter, eds.) 53–71. Springer, Berlin.

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

RAFAJŁOWICZ, E. and MYSZKA, W. (1988). Optimum experimental design for a regression on a hypercube - generalization of Hoel's result. *Ann. Inst. Statist. Math.* **40** 821–827.

SCHWABE, R. (1994). Optimal designs for additive linear models. *Statistics* (to appear).

SCHWABE, R. (1995). Experimental design for linear models with higher order interaction terms. In *Symposia Gaussiana. Proceedings of the 2nd Gauss Symposium, Conference B: Statistical Sciences, München 1993.* (H. Schneeweiß and V. Mammitzsch, eds.) 281–288. DeGruyter, Berlin.

SCHWABE, R. and WIERICH, W. (1995). D-optimal designs of experiments with non-interacting factors. *J. Statist. Plann. Inference* **44** 371–384.

FREIE UNIVERSITÄT BERLIN
1. MATHEMATISCHES INSTITUT
ARNIMALLEE 2–6
14 195 BERLIN
GERMANY