E-optimum Designs in Linear Models with Both Qualitative and Quantitative Factors of Influence

Rainer Schwabe*

Freie Universität Berlin, 1. Mathematisches Institut,
Arnimallee 2-6, 14195 Berlin, Germany

Abstract

In most experimental situations the outcome is influenced by various factors and different interaction structures. For a variety of simple but practically useful examples we investigate the generation of E-optimum designs, i. e. designs which minimize the maximum variance for all normalized estimates.

1 Introduction

We consider situations in which the outcome of an experiment depends on two or more factors of influence. In particular, the response function associated with the experiment can be linearly parametrized, i. e. the observation $Y(t) = \sum_{i=1}^{p} a_i(t)\beta_i + Z$ can be represented as a linear combination of the known regression functions $a_1, ..., a_p$ with unknown coefficients $\beta_1, ..., \beta_p$, where $t = (t_1, ..., t_K)'$ is the setting of the K factors of influence and Z denotes random noise.

For N different experiments which can be realized according to the linear model $Y_n(t^{(n)}) = a(t^{(n)})'\beta + Z_n$ the experimental condition $t^{(n)} \in T$ can be chosen by the experimenter according to the design $d = (t^{(1)}, ..., t^{(N)})$, where replicates may occur.

Under the usual assumptions of zero mean and uncorrelated equally dispersed errors the model is completely described by its response function $E(Y(t)) = a(t)'\beta$, $t \in T$. The quality of a design is measured by its information matrix $\mathbf{I}(d) = \frac{1}{N}\sum_{n=1}^{N} a(t^{(n)})a(t^{(n)})'$ because the covariance matrix $\mathrm{Cov}(\widehat{\beta}_d)$ of the least squares estimator $\widehat{\beta}_d$ is proportional to the inverse $\mathbf{I}(d)^{-1}$ of the information matrix (in case $\mathbf{I}(d)$ is regular).

To simplify the optimization problem we make use of the notion of (generalized) designs δ on T and the associated information matrices $\mathbf{I}(\delta) = \int_T aa'\, d\delta$. For a design d of size N consider the corresponding generalized design δ_d which assigns equal weight $\frac{1}{N}$ to $t^{(1)}, ..., t^{(N)}$. Then $\mathbf{I}(\delta_d) = \mathbf{I}(d)$ which shows that the generalization is straightforward. For further readings on the use of generalized designs we refer to Fedorov (1972), Kiefer (1974) or Bandemer et al. (1977).

Since, in general, it is impossible to find a design which maximizes the information matrix uniformly we have to investigate some real valued criterion function. In the present note we consider the minimax criterion of minimizing the maximum variance for all normalized estimates of linear combinations of the unknown parameters. By Rayleigh's formula this is equivalent to the formulation: A design δ^* is E-optimum if it minimizes $\lambda_{\max}(\mathbf{C}(\delta))$, where $\mathbf{C}(\delta)$ is the generalized covariance matrix, $\mathbf{C}(\delta_d) = \mathrm{Cov}(\widehat{\beta}_d)$, and $\lambda_{\max}$ denotes the largest eigenvalue. In case of a minimal parametrization, i. e. the regression

*research supported by grant Ku719/2-2 of the Deutsche Forschungsgemeinschaft

functions are linearly independent on T, the E-optimality can also be expressed as the maximization of the smallest eigenvalue $\lambda_{\min}(\mathbf{I}(\delta))$ of the information matrix. In contrast to other commonly used criteria the E-criterion is not differentiable, in general, and the construction of E-optimum designs is more difficult (see Kiefer, 1974).

In the present note we investigate situations in which product designs are E-optimum. Note that for finitely supported marginal designs δ_1 and δ_2 which assign weights w'_m and w''_n to experimental conditions $t_1^{(m)} \in T_1$ and $t_2^{(n)} \in T_2$, respectively, the product design $\delta_1 \otimes \delta_2$ is defined as that design which assigns weights $w_{m,n} = w'_m w''_n$ to each level combination $t^{(m,n)} = (t_1^{(m)}, t_2^{(n)}) \in T = T_1 \times T_2$.

In general, product designs are E-optimum in models with product type interaction structures and their marginals are the E-optimum designs in the corresponding single factor models (Section 2). Subsequently those models are investigated in which, at least, one of the factors is qualitative, i. e. this factor may be adjusted to, at most, finitely many different levels. By invariance considerations also there E-optimum product designs can be obtained. However, the E-optimality of the marginals is preserved only in some particular cases. Finally, some results are presented for a variety of more complicated interaction structures.

2 Complete product-type interaction

Given the single factor models $E(Y_1(t_1)) = f_1(t_1)'\beta^{(1)}$, $t_1 \in T_1$, and $E(Y_2(t_2)) = f_2(t_2)'\beta^{(2)}$, $t_2 \in T_2$, the general two-factor model with complete product-type interaction is defined by

$$E(Y(t_1,t_2)) = \textstyle\sum_{i=1}^{p_1} \sum_{j=1}^{p_2} f_{1i}(t_1) f_{2j}(t_2) \beta_{ij}, \quad t_1 \in T_1 \times T_2. \tag{1}$$

By means of the equivalence theorem for E-optimality (see Pukelsheim, 1993, or Ermakov and Melas, 1995) the following general result can be obtained (Schwabe, 1996, Theorem 4.10, see also Rafajłowicz and Myszka, 1988, for the case of simple maximum eigenvalues) which can immediately be extended to an arbitrary number of K different factors,

Theorem 1. *The product design $\delta^* = \delta_1^* \otimes \delta_2^*$ is E-optimum in the model with complete product-type interactions if the marginal designs δ_1^* and δ_2^* are E-optimum in the associated single factor models.*

3 One-way layout with an additional factor: no general mean

In the following we consider models in which one of the factors is qualitative and is described by a one-way layout $E(Y_1(i)) = \alpha_i$. For the two-factor additive model

$$E(Y(i,x)) = \alpha_i + f(x)'\beta, \quad i = 1, ..., I, \ x \in \mathcal{X} \tag{2}$$

the corresponding second marginal model $E(Y_2(x)) = \mu_0 + f(x)'\beta$ includes a constant term μ_0. By invariance considerations with respect to permutations of the treatment levels i of the first factor it can be seen that every design δ is dominated by its symmetrization

$\overline{\delta}_1 \otimes \delta_2$, where $\overline{\delta}_1$ is the uniform design on the levels of the first factor, which assigns equal weights $\frac{1}{I}$ to each level, and δ_2 is the second marginal of δ, $\delta_2(x) = \sum_{i=1}^{I} \delta(i,x)$ (see e. g. Schwabe, 1996, Subsection 6.2). It remains to find an optimum marginal design δ_2.

Theorem 2. *The product design $\delta^* = \overline{\delta}_1 \otimes \delta_2^*$ is E-optimum in the additive model without general mean if the marginal design δ_2^* is E-optimum in the transformed second marginal model $E(Y_2(x)) = \mu_0 + \sqrt{I} f(x)'\beta$, $x \in \mathcal{X}$.*

The proof is based on the fact that the eigenvalues of the information matrix for the two-factor model coincide with that for the transformed marginal model up to a factor $\frac{1}{I}$. For the particular case that in the second marginal model the E-optimum design δ_2^* is centered, i. e. $\int f \, d\delta_2^* = 0$, this design δ_2^* constitutes an E-optimum product design $\overline{\delta}_1 \otimes \delta_2^*$ (Schwabe, 1995).

Example 1. *Two-way layout:* $E(Y(i,j)) = \alpha_i + \beta_j$, $i = 1, ..., I$, $j = 1, ..., J$, with a control level J for the second factor, i. e. $\beta_J = 0$. Then the product design $\overline{\delta}_1 \otimes \delta_2^*$ is E-optimum which assigns equal weights $2/(I^2 + 4I(J-1))$ to all treatment combinations (i,j), $j \neq J$, and weights $(I + 2(J-1))/(I^2 + 4I(J-1))$ to the treatment-control combinations (i, J).

4 One-way layout with an additional factor: with general mean

The two-factor additive model of the previous section can also be written as

$$E(Y(i,x)) = \mu_0 + \alpha_i + f(x)'\beta, \quad i = 1, ..., I, \; x \in \mathcal{X} \tag{3}$$

with a general mean μ_0. The identifiability of the parameters will be ensured by imposing an additional condition $\sum_{i=1}^{I} \alpha_i = 0$ such that the parameters α_i now represent the effects of level i compared to their mean. Again, by invariance considerations every design δ is dominated by its symmetrization $\overline{\delta}_1 \otimes \delta_2$ (Schwabe, 1996, Example 6.9) and it remains to find an optimum marginal design δ_2.

Theorem 3. *The product design $\delta^* = \overline{\delta}_1 \otimes \delta_2^*$ is E-optimum in the additive model with general mean if the marginal design δ_2^* is E-optimum in the second single factor model.*

The proof is based on the block diagonal structure of the generalized covariance matrix where the blocks are related to the single factors.

5 One-way layout combined with regression: various interactions

In this section we confine to the broad class of models in which one of the factors is qualitative and follows a one-way layout and the second factor is described by a polynomial regression of degrees 1 or 2 (linear or quadratic). Also in this setting every design δ is dominated by its symmetrization $\overline{\delta}_1 \otimes \delta_2$ in case of a general symmetric parametrization

$$E(Y(i,x)) = f_0(x)'\beta^{(0)} + f_1(x)'\beta^{(i)}, \quad i = 1, ..., I, \; x \in \mathcal{X}. \tag{4}$$

In regression models usually an interval $\mathcal{X} = [b_1, b_2]$ is assumed for the marginal design region.

Example 2. *One-way layout combined with linear regression:* All possible kinds of interaction of one qualitative factor with additional linear regression are listed in Table 1. In all these situations there is a (product) design $\overline{\delta}_1 \otimes \delta_2^*$ which is E-optimum. By a majorization argument the optimum marginal design δ_2^* is concentrated on the two endpoints b_1 and b_2.

In case of a symmetric interval $\mathcal{X} = [-b, b]$ the E-optimum design δ_2^* in the marginal model which assigns equal weights $\frac{1}{2}$ to both endpoints yields a block diagonal structure of the information matrix $\mathbf{I}(\overline{\delta}_1 \otimes \delta_2^*)$ and, hence, the resulting design $\overline{\delta}_1 \otimes \delta_2^*$ is E-optimum for all of the present models.

$E(Y(i,x))$	w_0^*	$\lambda_{\max}$
$\beta_0 + \beta_1 x$	$\dfrac{b^2+2}{b^2+4}$	$1 + \dfrac{4}{b^2}$
$\beta_{0i} + \beta_1 x$	$\dfrac{Ib^2+2}{Ib^2+4}$	$I + \dfrac{4}{b^2}$
$\beta_0 + \beta_{1i} x$	$\dfrac{b^2+2I}{b^2+4I}$	$1 + \dfrac{4}{b^2} I$
$\beta_{0i} + \beta_{1i} x$	$\dfrac{b^2+2}{b^2+4}$	$I + \dfrac{4}{b^2} I$

Table 1: One qualitative factor combined with linear regression, optimum weights

For an interval $\mathcal{X} = [0, b]$ the optimum weights w_0^* are given in Table 1, together with the corresponding maximum eigenvalues of the (generalized) covariance matrix for the resulting design $\overline{\delta}_1 \otimes \delta_2^*$, where δ_2^* assigns weight w_0^* to 0 and weight $1 - w_0^*$ to b. The last entry of Table 1 shows complete product-type interactions and the optimum weight is obtained by Theorem 1.

Example 3. *One-way layout combined with quadratic regression:* All possible interaction structures are exhibited in Table 2. The underlying marginal design region is the unit intervall $\mathcal{X} = [-1, 1]$. By a majorization argument we find that an E-optimum design $\overline{\delta}_1 \otimes \delta_2^*$ is given by a symmetric design δ_2^* which assigns weight w_0^* to the midpoint 0 and equal weights $\frac{1}{2}(1 - w_0^*)$ to both endpoints -1 and 1 of the interval. The optimum weights w_0^* are given in Table 2, together with the corresponding maximum eigenvalues of the covariance matrix for the resulting design $\overline{\delta}_1 \otimes \delta_2^*$. Different interaction structures give different optimum weights depending on the number I of levels of the qualitative factor as can be seen from the table. In particular, if I is large these differences are substantial.

$E(Y(i,x))$	w_0^*	$\lambda_{\max}$
$\beta_0 + \beta_1 x + \beta_2 x^2$	$\dfrac{3}{5}$	5
$\beta_{0i} + \beta_1 x + \beta_2 x^2$	$\dfrac{I+2}{I+4}$	$I + 4$
$\beta_0 + \beta_{1i} x + \beta_2 x^2$	$\dfrac{2I-1}{I^2+I-1}$	$I + 2 + \dfrac{1}{I-1}$
$\beta_0 + \beta_1 x + \beta_{2i} x^2$	$\dfrac{2I+1}{4I+1}$	$4I + 1$
$\beta_{0i} + \beta_{1i} x + \beta_2 x^2$	$\dfrac{I}{2I-1}$	$2I + 1 + \dfrac{1}{I-1}$
$\beta_{0i} + \beta_1 x + \beta_{2i} x^2$	$\dfrac{3}{5}$	$5I$
$\beta_0 + \beta_{1i} x + \beta_{2i} x^2$	$\dfrac{2I+1}{4I+1}$	$4I + 1$
$\beta_{0i} + \beta_{1i} x + \beta_{2i} x^2$	$\dfrac{3}{5}$	$5I$

Table 2: One qualitative factor combined with quadratic regression, optimum weights

A different structure of interactions is exhibited by the intra-class experiments

$$E(Y(i,x)) = f_i(x)'\beta^{(i)}, \quad x \in \mathcal{X}_i, \ i = 1, ..., I, \tag{5}$$

in which the regression functions $f_i = (f_{i1}, ..., f_{ip_i})'$ and the conditional design region $\mathcal{X}_i$ may also depend on the level i of the qualitative factor. With the notion of conditional designs (see Kurotschka, 1984) the E-optimum designs can readily be characterized (Schwabe, 1996, Theorem 7.1),

Theorem 4. *Let, for every i, the design $\delta^*_{2|1}(i,\cdot)$ be E-optimum in the conditional model $E(Y_i(x)) = f_i(x)'\beta_i$, $x \in \mathcal{X}_i$, associated with the level i and let δ^*_1 assign weights $\lambda_i/(\sum_{j=1}^I \lambda_j)$ to level i, where λ_i is the largest eigenvalue of the conditional (generalized) covariance matrix for $\delta^*_{2|1}$. Then the mixture $\delta^* = \delta^*_1 \otimes \delta^*_{2|1}$ is E-optimum in the intra-class model.*

6 Concluding remarks

The methods proposed throughout the paper permit a reduction of the problem to determine optimum designs for multi-factor models to components corresponding to single factors. Since a lot of optimum designs are already known for one-factor models and optimization is much easier in one dimension this enables us to construct optimum designs for complicated models which would be difficult to treat directly. For example, it would be interesting to transfer the known results on E-optimum designs in polynomial models (see e. g. Heiligers, 1994) to the present settings with an additional one-way layout.

Related results have been obtained by Kurotschka, Schwabe and Wierich (1992) and Schwabe (1996) on D- and A-optimum designs aiming at minimizing the determinant and the trace of the (generalized) covariance matrix, respectively.

Finally, we want to emphasize once more that the structure of the interactions has a strong influence on the performance of a design.

References

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

ERMAKOV, S. M. and V. B. MELAS (1995). *Design and Analysis of Simulation Experiments.* Kluwer, Dordrecht.

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

HEILIGERS, B. (1994). E-optimal design in weighted polynomial regression. *Ann. Statist.* **22**, 917–929.

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

KUROTSCHKA, V. G. (1984). A general approach to optimum design of experiments with qualitative and quantitative factors. In J. K. GHOSH and J. ROY (eds.), *Statistics: Applications and New Directions: Proceedings of the Indian Statistical Institute Golden Jubilee International Conference Calcutta 1981*, 353–368.

KUROTSCHKA, V., R. SCHWABE and W. WIERICH (1992). Optimum designs for partly interacting qualitative and quantitative factors. In A. PÁZMAN and J. VOLAUFOVÁ (eds.), *Proba-*

stat '91, Proc. Int. Conf. Probab. Math. Statist., Bratislava 1991. Slovak Academy of Sciences, Bratislava, 102–108.

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

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

SCHWABE, R. (1995). Optimal designs for additive linear models. *Statistics* (to appear).

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