Ann. Inst. Statist. Math.
Vol. 51, No. 4, 723–730 (1999)

EFFICIENCY BOUNDS FOR PRODUCT DESIGNS IN LINEAR MODELS

RAINER SCHWABE[1]* AND WENG KEE WONG[2]**

[1] *Fachbereich Mathematik, Universität Mainz, 55099 Mainz, Germany*
[2] *Department of Biostatistics, University of California, Los Angeles, 10833 Le Conte Ave., Los Angeles, CA 90095, U.S.A.*

(Received February 23, 1998; revised June 26, 1998)

Abstract. We provide lower efficiency bounds for the best product design for an additive multifactor linear model. The A-optimality criterion is used to demonstrate that our bounds are better than the conventional bounds. Applications to other criteria, such as IMSE (integrated mean squared error) criterion are also indicated. In all the cases, the best product design appears to perform better when there are more levels in each factor but decreases when more factors are included. Explicit efficiency formulas for non-additive models are also constructed.

Key words and phrases: A-optimal designs, additive models, continuous designs, interaction, k-way layout, marginal models.

1. Introduction

Finding optimal or efficient designs is usually a difficult task in multi-dimensional models when different factors have to be incorporated. Simplifications can be expected if the model structure can be partitioned into components related to the influence of individual factors. In this paper, we are concerned with additive model structures of the form

$$Ey(t) = \beta_0 + \beta^T f(t) = \beta_0 + \sum_{i=1}^{k} \beta_i^T f_i(t_i), \qquad t \in T, \tag{1.1}$$

where $t = (t_1, \dots, t_k)^T$, $f(t)^T = (f_1(t_1)^T, \dots, f_k(t_k)^T)$ is the vector of regression functions on the design region T and $\beta^T = (\beta_1^T, \dots, \beta_k^T)$. The absence of mutual

* Now at Institut für Mathematik I, Freie Universität, Berlin, Arnimallee 2-6, 14195 Berlin, Germany. Essential parts of this work were done while Schwabe was visiting the University of Technology at Darmstadt. His work was partly supported by the research grant Ku719/2-2 of the DFG.

** The research of Wong was partially supported by a NIH research grant R29 AR44177-01A1.

influence among the factors leads to the natural assumption that $T = \times_{i=1}^{k} T_i$ is the product set of possible experimental settings $t \in T$ in which the k factor levels $t_1, \ldots, t_k$ may be chosen independently of each other. This model is popular because the effects of the individual factors are additive and no interactions are permitted among factors, see Clark (1965), Cook and Thibodeau (1980), and Wong (1994) where they assumed the factor levels are all continuous. This type of model is also common in the analysis of variance setup where each factor ranges over a finite set of qualitative levels. The latter situation occurs for instance in the standard k-way layout. Schwabe (1996, 1998) considers design issues for models where both qualitative and continuous factors are present.

Let p_i be the dimension of the regression functions f_i associated with the i-th factor, and assume that the components of each f_i are linearly independent, $i = 1, \ldots, k$. Additionally, we assume that the model is homoscedastic and that all observations are uncorrelated.

In practice, each factor in a multi-factor experiment may affect the mean response differently. To accommodate for this, we consider the marginal models given by

$$Ey(t_i) = \beta_{0,i} + \beta_i^T f_i(t_i), \qquad t_i \in T_i, \qquad i = 1, \ldots, k,$$

where β_i is a p_i-dimensional vector of parameters. These marginal models and their optimal designs will play a crucial role in the construction of efficient designs for the model (1.1).

All designs here are continuous designs in the sense that they are treated as finitely supported probability measures on the set T (Kiefer (1959)). Our focus is on product designs which have the form $\xi = \xi_1 \times \cdots \times \xi_k$ and each component ξ_i is a design defined on the marginal design region T_i, $i = 1, \ldots, k$. These product designs are easy to construct since they are pieced together from the designs from the 'smaller' marginal models. Because of their simplicity, these designs are perhaps the most widely used in multi-factor experiments, see for example, Rafajlowicz and Myszka (1988, 1992) and Wong (1994). An increasingly popular design strategy is to restrict the search for the optimal designs within the class of product designs (Schwabe (1996), Dette and Röder (1996)). The resulting optimal design is called the optimal product design and we will show in this paper that this type of design is generally rather efficient.

If we let $F(t)^T = (1, f(t)^T)$, the normalized information matrix of ξ relative to the full model in (1.1) is given by $M(\xi) = \int_T F(t)F(t)^T \xi(dt)$. Likewise, if $F_i(t_i)^T = (1, f_i(t_i)^T)$, the normalized information matrix of the marginal design ξ_i is $M_i(\xi_i) = \int_{T_i} F_i(t_i)F_i(t_i)^T \xi_i(dt_i)$. These matrices are used to gauge the worth of a design. For instance, for A-optimality, we seek to minimize the trace $\operatorname{tr} M(\xi)^{-1}$, and the A-efficiency of an arbitrary design ξ is given by $\operatorname{eff}_A(\xi) = \operatorname{tr} M(\xi_A)^{-1} / \operatorname{tr} M(\xi)^{-1}$ where ξ_A is an A-optimal design (see e.g. Pázman (1986), p. 82). This quantity is between 0 and 1 and its reciprocal denotes the number of times the design ξ has to be replicated in order to do as well as the optimal design. Thus designs with high efficiencies are sought in practice.

The ease of finding the optimal design depends heavily on whether the criterion is compatible with the product structure. Schwabe (1996, 1998) discussed this issue in detail and noted that D- and G-optimal designs can be obtained as the best product designs. Schwabe and Wong (1997) derived formulas for expressing the D- and G-efficiencies of product designs in terms of the D- and G-efficiencies of the marginal designs for suitable classes of multifactor linear models with hierarchical interaction structure. Analogous results are not available for other criteria such as A-optimality. Part of the problem is that the optimal product design may not be optimal among all designs.

The aim of the paper is to exploit the unique structure in additive linear models and provide sharper efficiency bounds for product designs. We use A-optimality as an illustrative criterion in the k-way layout but the method is applicable to other design criteria as well. For example, we state corresponding results for the integrated mean-squared error criterion commonly used in regression settings where the factors are continuous. Extensions to models with hierarchical interaction structures are briefly discussed in the last section.

2. A-optimal designs for additive k-way layouts

We present A-optimal designs when all factors are each at 2 levels including a control level for each factor. The marginal regression function here is $f_i(t_i) = t_i$ for $t_i = 0$ (control) and $t_i = 1$ (treatment), respectively. The aim is to show that finding the optimal designs can be laborious even in quite simple situations and so it is useful to have approximations for the optimal designs. The optimal designs found here will be used in Section 3 to assess the precision of our proposed lower efficiency bounds for the optimal product designs as approximations.

Table 1. Optimal moments a^* and b^* and A-efficiency of the best product design ξ^*.

k	a^*	b^*	$\operatorname{tr} M(\xi_A)^{-1}$	$(\operatorname{tr} M(\xi^*)^{-1})$	A-efficiency of ξ^*
2	0.4189	0.1937	10.6038	(10.6569)	0.995
3	0.4229	0.1969	15.3340	(15.4853)	0.990
4	0.4264	0.1997	20.0250	(20.3137)	0.985
5	0.4295	0.2021	24.6822	(25.1421)	0.981
10	0.4409	0.2108	47.5821	(49.2843)	0.965
100	0.4765	0.2363	433.505	(483.843)	0.895

Under this setting, the model equation is not affected by permutations of the factors, and an A-optimal design can be found in the class of symmetric designs ξ^s which are invariant with respect to these permutations. The information matrix of ξ^s is

$$M(\xi^s) = \begin{pmatrix} 1 & a\mathbf{1}_k^T \\ a\mathbf{1}_k & (a-b)\boldsymbol{I}_k + b\mathbf{1}_k\mathbf{1}_k^T \end{pmatrix}$$

where $\boldsymbol{I}_k$ and $\mathbf{1}_k$ denote the $k \times k$ identity matrix and the vector of all 1's in R^k, respectively.

The desired A-optimal design ξ_A is given by the values of a^* and b^* which minimize $\operatorname{tr} M(\xi^s)^{-1}$. These values are displayed in Table 1 together with $\operatorname{tr} M(\xi_A)^{-1}$ for selected values of k. The numbers in parentheses correspond to $\operatorname{tr} M(\xi^*)^{-1}$ for the best values of the first two moments for the product $\xi^* = \xi_1^* \times \cdots \times \xi_k^*$ of the one-dimensional marginal A-optimal designs ξ_i^* (i.e. for $a = \sqrt{2} - 1 = 0.4142$ and $b = a^2 = 0.1716$). The A-efficiency is obtained by dividing the two traces corresponding to ξ_A and ξ^*.

3. Lower bounds for the A-efficiency

We now discuss lower A-efficiency bounds useful for assessing the proximity of any design to the optimal one. These bounds are especially important in situations where the optimal design is unknown or is too laborious to compute. High efficiencies indicate that the performance of the design is close to the optimal.

3.1 *Conventional lower efficiency bounds*

Standard lower efficiency bound can be obtained from convex analysis arguments, see Pázman ((1986), p. 118) for example. For the A-criterion, one obtains

$$\operatorname{eff}_A(\xi) \geq 2 - \sup_{t \in T} F(t)^T M(\xi)^{-2} F(t) / \operatorname{tr} M(\xi)^{-1}$$

based on the directional derivative of the convex functional $\operatorname{tr} M(\xi)^{-1}$.

This bound may become negative if the number of factors k becomes large. As an example, consider identical marginal models $f_i = f_1$, $T_i = T_1$, and hence $\xi_i^* = \xi_1^*$. Then $\operatorname{eff}_A(\xi_1^* \times \cdots \times \xi_1^*) \geq 1 - ck(k-1)/\{k \operatorname{tr} M_1(\xi_1^*)^{-1} - (k-1)\}$, where c is a nonnegative constant which depends on the marginal model. In particular, for the k-way layout with equal numbers of levels for each factor (i.e. $p_i = p$) compared to the control, the efficiency bounds of the best product design are displayed in Table 2.

Here the regression function $f_i = (f_{i1}, \ldots, f_{ip})^T$ is given by $f_{ij}(t_i) = 1$ for $t_i = j$ and $f_{ij}(t_i) = 0$ otherwise, $t_i \in T_i = \{0, 1, 2, \ldots, p\}$. Similar results have been observed by Wierich (1989) under a different parametrization.

These bounds are not satisfactory because they may yield negative numbers and are rather inefficient, when compared with the actual efficiencies (Table 1). An alternative bound has been developed by Dette (1996) who proposed the lower efficiency bound

$$\operatorname{eff}_A(\xi) \geq \operatorname{tr} M(\xi)^{-1} / \sup_{t \in T} F(t)^T M(\xi)^{-2} F(t).$$

He based his arguments on the directional derivative of the concave functional $\{\operatorname{tr} M(\xi)^{-1}\}^{-1}$ instead of the convex functional $\operatorname{tr} M(\xi)^{-1}$. For the k-way layout setup when all factors have the same number of levels these bounds are presented in Table 3.

These bounds are better than those provided in Table 2. However, for large k these efficiency bounds tend to zero, suggesting a bad performance when a large number of factors is involved.

Table 2. Lower A-efficiency bounds for the best product design; convex optimization (k-way layout with $p_1 = \cdots = p_k = p$; "—" denotes negative values).

$p\backslash k$	2	3	5	6	10
1	0.81	0.61	0.20	—	—
2	0.90	0.79	0.59	0.49	0.07
3	0.90	0.81	0.62	0.53	0.15
4	0.91	0.83	0.66	0.57	0.24

Table 3. Lower A-efficiency bounds for the best product design; concave optimization (k-way layout with $p_1 = \cdots = p_k = p$).

$p\backslash k$	2	3	5	100
1	0.84	0.72	0.55	0.04
2	0.90	0.83	0.71	0.08
3	0.91	0.84	0.72	0.09
4	0.92	0.85	0.74	0.10

3.2 *Lower efficiency bounds based on the additive structure*

To improve the lower efficiency bound, we consider the A-optimal designs ξ_{β_i} for estimating the direct effect β_i in the marginal model. This means ξ_{β_i} minimizes $\operatorname{tr} C_i(\xi_i)^{-1}$ where $C_i(\xi_i) = \int_{T_i} f_i(t_i) f_i(t_i)^T \xi_i(dt_i) - \int_{T_i} f_i(t_i)\xi_i(dt_i) \int_{T_i} f_i(t_i)^T \xi_i(dt_i)$. If we use $\xi_1^* \times \xi_{\beta_2} \times \cdots \times \xi_{\beta_k}$ as a reference design within the set of product designs, we obtain for every arbitrary design ξ by a refinement argument

$$\operatorname{tr} M(\xi)^{-1} \geq \operatorname{tr} M_1(\xi_1)^{-1} + \sum_{i=2}^{k} \operatorname{tr} C_i(\xi_i)^{-1} \geq \operatorname{tr} M_1(\xi_1^*)^{-1} + \sum_{i=2}^{k} \operatorname{tr} C_i(\xi_{\beta_i})^{-1}$$

where ξ_i is the projection of ξ onto the i-th component. Hence, by replacing the first factor by any other we establish

THEOREM 1. *Let ξ_i^* be A-optimal in the i-th marginal model, $i = 1, \ldots, k$. Then $\xi_1^* \times \cdots \times \xi_k^*$ is an A-optimal product design in the additive model* (1.1) *and*

$$\begin{aligned}&\operatorname{eff}_A(\xi_1^* \times \cdots \times \xi_k^*)\\ &\quad \geq \frac{\sum_{i=1}^{k} \operatorname{tr} C_i(\xi_{\beta_i})^{-1} + \max_{1\leq i\leq k}\{\operatorname{tr} M_i(\xi_i^*)^{-1} - \operatorname{tr} C_i(\xi_{\beta_i})^{-1}\}}{\sum_{i=1}^{k} \operatorname{tr} M_i(\xi_i^*)^{-1} - (k-1)} > 0.\end{aligned}$$

For the additive k-way layout, when all factors have the same number of levels p, this lower bound simplifies to

$$\operatorname{eff}_A(\xi_1^* \times \cdots \times \xi_k^*) \geq 1 - 2(k-1)p(\sqrt{p+1} - \sqrt{p})/\{k(p + \sqrt{p+1})^2 - k + 1\}$$

Table 4. Lower A-efficiency bounds for the best product designs, structural approach (k-way layout with $p_1 = \cdots = p_k = p$).

$p \backslash k$	2	3	5	$\to \infty$
1	0.92	0.89	0.86	0.82
2	0.95	0.93	0.92	0.90
3	0.96	0.95	0.94	0.93
4	0.97	0.96	0.96	0.95

and the corresponding values are given in Table 4. The last column indicates the limiting behavior when k tends to infinity. A comparison with the convential bounds in Table 2 and Table 3 shows that these new bounds are substantially better, in particular for large k. Moreover, it becomes evident that the best product designs have high efficiencies even when the number of factors is large.

In regression settings with continuous level factors, A-optimality is not appropriate. A more suitable criterion which has properties similar to A-optimality is the integrated mean-squared error criterion defined by $\mathrm{IMSE}(\xi) = \int_T \mathrm{var}(\hat{\beta}_0 + \hat{\beta}^T f(t))\lambda(dt) = \mathrm{tr}\, M(\xi)^{-1} M(\lambda)$ where λ is uniform on the design region T. The IMSE can be identified as an A-criterion if the regression functions are chosen to be orthonormal with respect to the uniform measure λ. Hence, analogous efficiency bounds can be obtained when the A-criterion $\mathrm{tr}\, C_i(\xi_i)^{-1}$ concerning the direct effect β_i is replaced by the integrated mean squared error $\mathrm{tr}\, C_i(\xi_i)^{-1} C_i(\lambda_i)$ adjusted for the general mean $\int_{T_i}(\beta_0 + \beta_i^T f_i(t_i))\lambda_i(dt_i)$, where λ_i is uniform on T_i. Using, this observation Schwabe and Wong (1998) demonstrated that the efficiency bounds for the IMSE-optimal product designs are also better than those obtained from the conventional bounds, for an additive quadratic regression model. In general, the methodology proposed here can be used to obtain efficiency bounds for the Φ_q-criteria based on the eigenvalues of the information matrix and for criteria based on the variance $F(t)^T M(\xi)^{-1} F(t)$ of the predicted response.

4. Extensions to hierarchical interaction structure

We conclude this paper by mentioning that the ideas in Section 3 can be applied to construct lower efficiency bounds for more complicated models as long as the models have hierarchical interaction structures. The D-optimality of the best product designs for such models has been treated by Schwabe (1998). Specifically, let $\bigotimes$ denote the usual Kronecker product, and consider the multifactor model given by

$$Ey(t_1, \ldots, t_k) = \sum_{\delta \in \Delta} \bigotimes_{i \in \delta} f_i(t_i)^T \beta_\delta$$

where Δ is a subset of the power set of $\{1, \ldots, k\}$ and $\bigotimes_{i \in \emptyset} = 1$. For example, if $\Delta = \{\emptyset, \{1\}, \ldots, \{k\}\}$, we have the additive model; if Δ is the complete power set we have the Kronecker-product model and if $\Delta = \{\emptyset, \{1\}, \ldots, \{k\}, \{1,2\}, \ldots, \{k-1, k\}\}$, we have the first-order interaction model.

Table 5. A-efficiency bounds of the best product design for the first-order interaction model.

$p\backslash k$	3	5	$\to\infty$
1	0.81	0.73	0.68
2	0.88	0.83	0.81
3	0.91	0.88	0.87
5	0.94	0.93	0.92
10	0.97	0.96	0.96

We say the model has hierarchical interaction structure if $\delta \in \Delta$ implies that all corresponding lower order interactions are also included in the model, i.e. $\delta' \in \Delta$ for all subsets δ' of $\delta \in \Delta$. Examples of models with such structure are the three models just mentioned. Product designs for these models satisfy $\operatorname{tr} M(\xi_1 \times \cdots \times \xi_k)^{-1} = \sum_{\delta\in\Delta} \prod_{i\in\delta}\{\operatorname{tr} M_i(\xi_i)^{-1} - 1\}$. Thus, the product of the A-optimal marginal designs is the best product design.

In particular, for the best product design in the k-way layout with equal numbers of treatments ($p_i = p$) and first-order interactions we have

$$\begin{aligned}&\operatorname{eff}_A(\xi_1^* \times \cdots \times \xi_k^*)\\ &\quad\geq \frac{(p+\sqrt{p+1})^4 + (k-2)(p+\sqrt{p})^2 + \frac{1}{2}(k+1)(k-2)(p+\sqrt{p})^4}{\frac{1}{2}k(k-1)(p+\sqrt{p+1})^4 - k(k-2)(p+\sqrt{p+1})^2 + \frac{1}{2}(k-1)(k-2)}.\end{aligned}$$

Selected numerical results are presented in Table 5. Due to the complexity of the models these efficiency bounds are lower than the corresponding results in Table 4. When $k = 2$, it can be shown that the best product design is A-optimal (Rafajlowicz and Myszka (1988)), and, hence, its efficiency is equal to one. For $k > 2$, the present bounds are useful, because we cannot determine the A-optimal design explicitly and the actual A-efficiencies of the best product design are thus unknown.

Also in the present situation of first order interactions the efficiency bounds based on the model structure are substantially better than those obtained by pure convexity arguments, in particular, for large k.

References

Clark, V. (1965). Choice of levels in polynomial regression with one or two variables, *Technometrics*, **7**, 325–333.

Cook, R. D. and Thibodeau, L. A. (1980). Marginally restricted D-optimal designs, *J. Amer. Statist. Assoc.*, **75**, 366–371.

Dette, H. (1996). Lower bounds for efficiencies with applications, *Research Developments in Probability and Statistics, Festschrift in Honor of Madan Puri on the Occasion of his 65th Birthday* (eds. E. Brunner and M. Denker), 111–124, VSP, Utrecht.

Dette, H. and Röder, I. (1996). Optimal product designs for multivariate regression with missing terms, *Scand. J. Statist.*, **23**, 195–208.

Kiefer, J. (1959). Optimum experimental designs, *J. Roy. Statist. Soc. Ser. B*, **21**, 272–319.

Pázman, A. (1986). *Foundations of Optimum Experimental Design*, D. Reidel, Dordrecht.

Rafajlowicz, 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.

Rafajlowicz, E. and Myszka, W. (1992). When product type experimental design is optimal? Brief survey and new results, *Metrika*, **39**, 321–333.

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

Schwabe, R. (1998). Optimal designs for hierarchical interaction structures, *J. Statist. Plann. Inference*, **70**, 181–190.

Schwabe, R. and Wong, W. K. (1997). A relationship between efficiencies of marginal designs and product designs, *Metrika*, **45**, 253–257.

Schwabe, R. and Wong, W. K. (1998). Improved lower efficiency bounds for product designs in multifactor experiments (extended version), Berichte zur Stochastik und verwandten Gebieten, No. 98-4, Johannes Gutenberg-Universität Mainz.

Wierich, W. (1989). On the relative efficiency of the best second order proportional ANOVA designs, *Comput. Statist. Data Anal.*, **8**, 119–134.

Wong, W. K. (1994). Multifactor *G*-optimal designs with heteroscedastic errors, *J. Statist. Plann. Inference*, **40**, 127–133.