Open access • Journal Article • DOI:10.1080/01621459.1980.10477478

# Marginally restricted D-optimal designs — Source link

R. Dennis Cook, L. A. Thibodeau

**Institutions:** University of Minnesota, Harvard University

**Published on:** 01 Jun 1980 - Journal of the American Statistical Association (Taylor & Francis Group)

**Topics:** Optimal design and Equivalence (measure theory)

# Marginally Restricted

# D-optimal Designs

by

R. Dennis Cook and L. A. Thibodeau

Technical Report No. 331

Department of Applied Statistics
University of Minnesota
Saint Paul, Minnesota 55108

November 1, 1978

This work was supported in part by Grant #1-R01-GM25587-01 from the National Institute of General Medical Science.

## Abstract

In experimental design it often happens that some of the relevant carriers cannot be specified by the experimenter. We consider the problem of obtaining approximate D-optimal designs when the design space is a product space and the carriers associated with one margin are not subject to control. An equivalence theorem for D-optimal designs is presented. The essential ingredients of iterative schemes for generating designs are discussed.

## 1. Introduction

In classical optimal design for regression models it is usually assumed that all relevant carriers (independent variables) can be controlled completely throughout the design space, (see, for example, Fedorov, 1972). However, in many areas of application it is common to find that some of the known carriers are not subject to control. This often happens when the experiment consists of applying levels of a "treatment" to experimental units which differ on known relevant quantitative variables. In this case the values of the carriers associated with the experimental units are restricted by the availability of the units.

Harville (1974, 1975) discusses the problem of obtaining nearly optimal allocation of experimental units for analysis of covariance models. He presents algorithms which result in nearly D-optimal exact designs for inferences about the treatment effects in additive covariance models and discusses extensions to nonadditive models.

Here, we consider the problem of obtaining D-optimal designs for regression models when the values of some of the carriers are restricted and not subject to control by design. We first present the general formulation and some relevant background information.

Let $\underline{f}'(x) = (f_1,\ldots,f_p)$ denote a vector of $p$ linearly independent continuous functions on some compact space $\chi$. An experiment consists of selecting an $x$ in $\chi$ and observing a random variable $y(x)$ with regression function $E(y|x) = \underline{f}'\underline{\theta}$ and constant variance $\sigma^2$. We assume that the $f_i$ are known while the parameter vector $\underline{\theta}$ is unknown. If $\xi$ is a probability measure on $\chi$ then $\xi$ defines an experimental design.

Exact designs concentrate mass $\xi(x_i)$ at points $x_i$, $i=1,2,\ldots,r$, subject to the restriction that $N\xi(x_i) = n_i$ is integral for all $i$. An exact design specifies that the experimenter is to take $N$ uncorrelated observations, $n_i$ at $x_i$. The resulting covariance matrix of the least squares estimate of $\underline{\theta}$ is of the form

$$\left(\frac{\sigma^2}{N}\right) \underline{M}^{-1}(\xi)$$

where the information matrix, $\underline{M}(\xi)$, is

$$\underline{M}(\xi) = \int_{\mathcal{X}} \underline{f}\, \underline{f}'\, d\xi \quad .$$

Approximate designs are not constrained by the requirement that $n_i$ be integral for all $i$. Here we consider only approximate designs.

The choice of a design is often based on the minimization of some functional of the information matrix, $\underline{M}(\xi)$. Perhaps the two most commonly employed functionals are

(i) determinant $\underline{M}^{-1}(\xi) = |\underline{M}^{-1}(\xi)|$

and

(ii) $\max\limits_{x \varepsilon \xi} d(x;\xi)$

where $d(x;\xi) = \underline{f}'(x)\underline{M}^{-1}(\xi)\underline{f}(x)$. Designs minimizing these functionals are called D and G-optimal designs, respectively. The following result due to Kiefer and Wolfowitz (1960) established the equivalence of D and G-optimal approximate designs and provided a way of verifying whether a given design is D-optimal:

Theorem 1: (Equivalence Theorem). The following conditions are equivalent.

(i) $|M^{-1}(\xi_D)| = \min_{\xi} |M^{-1}(\xi)|$

(ii) $\max_{x} d(x;\xi_D) = \min_{\xi} \max_{x} d(x;\xi)$

(iii) $\max_{x} d(x;\xi_D) = p$ .

The set of all designs satisfying these conditions is convex and the corresponding information matrices are identical.

In the next section we provide analogous equivalences for situations in which $x$ is a vector, $x = (x_1,x_2)$, and the values of $x_1$ to be included in the experiment are not at the experimenter's control.

## 2. Marginally Restricted D-Optimal Designs

Let $x = (x_1, x_2)$ and let $\xi(x) = \xi(x_1, x_2)$ denote an arbitrary design on $\chi = \chi_1 \times \chi_2$. We consider only designs for which $|\underline{M}(\xi)| \neq 0$. Let $\xi_i$, $i=1,2$, denote the marginal design

$$\xi_i(x_i) \equiv \int_{\chi_j} d\xi(x_1, x_2) \quad , i \neq j = 1, 2 .$$

Since the values of $x_1$ to be included in the experiment are not subject to control, we assume that they specify a marginal design $\xi_1^*$, say, which places mass at points of a finite collection $S_1^*$. Following Fedorov (1972), we refer to $S_1^*$ as the spectrum of the design $\xi_1^*$. All permissible designs must have $\xi_1 = \xi_1^*$ and we assume that there is at least one permissible design $\xi$ such that $|\underline{M}(\xi)| \neq 0$.

Let $C = \{\xi | \xi_1 = \xi_1^*\}$ and note that $C$ is convex. The associated family of information matrices, $\{\underline{M}(\xi) | \xi \epsilon C\}$, has the same properties as the family of all information matrices (cf, Fedorov, 1972, p. 66). In particular, under the assumption that $S_1^*$ is finite we may, without loss of generality, restrict $C$ to measures with finite spectrum. Let $S_2$ denote the spectrum of $\xi_2$

The design problem is to choose the "best" design from $C$ according to the following definition.

<u>Definition 1</u>: The design $\hat{\xi}$ is a marginally restricted D-optimal design if

$$\min_{\xi \epsilon C} |\underline{M}^{-1}(\xi)| = |\underline{M}^{-1}(\hat{\xi})| .$$

In the case that $\underline{f}(x_1, x_2) = \underline{g}_1(x_1) \otimes \underline{g}_2(x_2)$, where $\otimes$ denotes the Kronecker product, a marginally restricted D-optimal design is easily determined.

Lemma 1: If $\underline{f}(x_1,x_2) = \underline{g}_1(x_1) \otimes \underline{g}_2(x_2)$ on $\chi = \chi_1 \times \chi_2$ then a marginally restricted D-optimal design is equal to $\xi_2^D \times \xi_1^*$ where $\xi_2^D$ is the D-optimal design for $\underline{g}_2$ on $\chi_2$.

Proof: The result follows immediately from Hoel (1965).

Recall that for any design

$$\int_{\chi} d(x_1,x_2;\, \xi)\, d\xi(x_1,x_2) = p$$

and, thus,

$$\max_{x_1,x_2} d(x_1,x_2;\, \xi) \geq p \quad .$$

The following lemma establishes an analogous result for the maximum over the unrestricted margin, $\chi_2$. First, for $\xi \varepsilon C$, let $\xi_{2|1}(x_2|x_1)$ denote the associated conditional design on $\chi_2$ given $x_1 \varepsilon S_1^*$:

$$\xi_{2|1}(x_2|x_1) = \xi(x_1,x_2)/\xi_1^*(x_1) \text{ for } \xi_1^*(x_1) > 0.$$

Lemma 2: For $\xi \varepsilon C$

$$\int_{\chi_1} \max_{x_2} d(x_1,x_2;\, \xi)\, d\xi_1^*(x_1) \geq p \; .$$

Proof: The result follows immediately from the relationship

$$\int_{\chi_2} d(x_1,x_2;\, \xi)\, d\xi_{2|1}(x_2|x_1) \leq \max_{x_2} d(x_1,x_2;\, \xi)$$

for all $x_1 \varepsilon S_1^*$.

The following theorem presents equivalences for marginally restricted D-optimal designs analogous to those of Theorem 1 for D-optimal designs.

Theorem 2: The following conditions are equivalent:

(i) $|M^{-1}(\hat{\xi})| = \min_{\xi \varepsilon C} |M^{-1}(\xi)|$

(ii) $\int_{\chi_1} \max_{x_2} d(x_1,x_2;\, \hat{\xi})\, d\xi_1^*(x_1)$

$$= \min_{\xi \varepsilon C} \int_{\chi_1} \max_{x_2} d(x_1,x_2;\ \xi)\ d\xi_1^*(x_1)$$

(iii) $$\int_{\chi_1} \max_{x_2} d(x_1,x_2;\hat{\xi})\ d\xi_1^*(x_1) = p\ .$$

The set of all designs satisfying these conditions is convex and the corresponding information matrices are identical.

Proof: The proof follows along the same lines as that for Theorem 1. Only the main points will be sketched here. We first show that (ii) and (iii) follow from (i).

Let $\hat{\xi}$ satisfy (i) and let $\xi$ denote an arbitrary design in C. Then $\xi_\alpha \equiv (1-\alpha)\hat{\xi} + \alpha\xi \ \varepsilon\ C$ for all $0 \leq \alpha \leq 1$. Since $|\underline{M}(\hat{\xi})| \geq |\underline{M}(\xi)|$ for all $\xi \varepsilon C$ we must have

$$\frac{\partial}{\partial \alpha} \log |\underline{M}(\xi_\alpha)|\Big|_{\alpha=0} \leq 0\ .$$

Or, after evaluation,

$$\int_{\chi_1} \int_{\chi_2} d(x_1,x_2;\ \hat{\xi})\quad d\xi_{2|1}(x_2|x_1)\ d\xi_1^*(x_1) \leq p$$

where $\xi_{2|1} = \xi(x_1,x_2)/\xi_1^*(x_1)$, $x_1 \ \varepsilon\ S_1^*$. Choosing, for each $x_1 \ \varepsilon\ S_1^*$, $\xi_{2|1}(x_2|x_1)$ to place mass 1 at the value of $x_2$ which achieves $\max_{x_2} d(x_1,x_2;\ \hat{\xi})$ it follows that

$$\int_{\chi_1} \max_{x_2} d(x_1,x_2;\ \hat{\xi})\ d\xi_1^*(x_1) \leq p\ .$$

This in combination with lemma 2 establishes results (ii) and (iii).

To show that (i) follows from (ii), let $\hat{\xi}$ satisfy (ii) and assume

$$|\underline{M}^{-1}(\hat{\xi})| > \min_{\xi \varepsilon C} |\underline{M}^{-1}(\xi)| .$$

There is a design $\xi \varepsilon C$ such that

$$\frac{\partial}{\partial\alpha} \log|\underline{M}(\xi_\alpha)|\Big|_{\alpha=0} = \int_{\chi} d(x_1,x_2;\ \hat{\xi})\ d\xi(x_1,x_2) - p > 0 .$$

However, since $\hat{\xi}$ satisfies (ii),

$$\int_{\chi} d(x_1,x_2;\ \hat{\xi})\ d\xi(x_1,x_2) \leq \int_{\chi_1} \max_{x_2} d(x_1,x_2;\ \hat{\xi})\ d\xi_1^*(x_1) \leq p$$

and result (i) follows. Other equivalences follow in a similar fashion.

As in the case of the equivalence theorem, Theorem 2 establishes equivalences between functionals based on the determinant and the variances of the predicted values, and provides conditions for verifying when a given design is the marginally restricted D-optimal design. However, the following necessary condition may be a bit easier to verify in practice.

Corollary 1: Let $\hat{\xi}$ denote a marginally restricted D-optimal design then

$$v = \int_{\chi_1} \max_{x_2 \varepsilon S_2} d(x_1,x_2;\ \hat{\xi})\ d\xi_1^*(x_1) = p .$$

Proof: By Theorem 2, $v \leq p$. Assume $v < p$, then

$$\int_{\chi_2} d(x_1,x_2;\ \hat{\xi})\ d\xi_{2|1}(x_2|x_1) \leq \max_{x_2 \varepsilon S_2} d(x_1,x_2;\ \hat{\xi})$$

and, thus,

$$p = \int_{\chi} d(x_1,x_2;\ \hat{\xi})\ d\hat{\xi}(x_1,x_2) \leq v < p .$$

The result follows by contradiction.

According to this corollary, to verify that a given design is not the marginally restricted design we need consider only the points in $S_2$.

Additive models represent an important special case which frequently occurs in practice. If the experimenter can specify that the model is additive and contains a constant term then $\underline{f}$ may be written as $\underline{f}'(x_1,x_2) = (1,\underline{g}_1'(x_1), \ \underline{g}_2'(x_2))$. The following lemma shows how to construct a marginally restricted D-optimal design in this case.

<u>Lemma 3</u>: If $\underline{f}'(x_1,x_2) = (1,\underline{g}_1'(x_1), \underline{g}_2'(x_2))$ then a marginally restricted D-optimal design is $\xi_2^D \times \xi_1^*$, where $\xi_2^D$ is the D-optimal design for $(1,\underline{g}_2')$ on $\chi_2$.

<u>Proof</u>: Let $\xi(x_1,x_2) = \xi_1(x_1) \times \xi_2(x_2)$,

$$\underline{M}_i = \int_{\chi_i} \tilde{\underline{g}}_i \tilde{\underline{g}}_i' \, d\xi_i(x_i)$$

and

$$d_i(x_i;\ \xi_i) = \tilde{\underline{g}}_i' \underline{M}_i^{-1} \tilde{\underline{g}}_i$$

where

$$\tilde{\underline{g}}_i' = (1,\underline{g}_i'(x_i)), \qquad i = 1,2 \ .$$

It is straightforward to verify that

$$d(x_1,x_2;\ \xi) = d_1(x_1;\ \xi_1) + d_2(x_2;\ \xi_2) - 1 \ .$$

The result follows immediately from Theorems 1 and 2 by setting $\xi_1 = \xi_1^*$ and $\xi_2 = \xi_2^D$.

In general, Lemma 3 will not hold for models without constant terms. This is easily seen by considering the case $\underline{f}'(x_1,x_2) = (x_1,x_2)$, $\chi = [-1,1]^2$. Also, it is worth noting that Lemma 3 shows how to obtain a D-optimal design for an additive model in the unrestricted case.

The following examples illustrate the use of Theorem 2 in two cases not covered by Lemmas 1 or 3.

Example 1: Let $\chi = [-1,1]^2$ and $\underline{f}'(x_1,x_2) = (1,x_2x_1,x_2)$. Also, let $\xi(x_1,x_2) = \xi_2^D(x_2) \times \xi_1^*(x_1)$ where $\xi_2^D$ is the D-optimal design for $(1,x_2)$ on [-1,1], i.e. $\xi_2^D(-1) = \xi_2^D(1) = 1/2$. It is easily verified that

$$d(x_1,x_2;\ \xi) = 1 + x_2^2 + x_2^2x_1^2/k$$

where

$$k = \int_{-1}^{1} x_1^2\ d\xi_1^*(x_1)\ .$$

Clearly,

$$\max_{x_2} d(x_1,x_2;\ \xi) = 2 + x_1^2/k$$

and

$$\int_{-1}^{1} (2+x_1^2/k)\ d\xi_1^* = 3\ .$$

Thus, by condition (iii) of Theorem 2, $\xi$ is a marginally restricted D-optimal design.

Example 2: Let $\chi = [-1,1]^2$ and $\underline{f}'(x_1,x_2) = (1,x_2x_1,x_2^2)$. Consider the design $\xi = \xi_2^D \times \xi_1^*$ where $\xi_2^D$ is the D-optimal design for $(1,x_2,x_2^2)$ on [-1,1], i.e. $\xi_2^D(-1) = \xi_2^D(0) = \xi_2^D(1) = 1/3$. A little algebra will verify that

$$d(x_1,x_2;\ \xi) = 3 - 6x_2^2 + \frac{3}{2k}\, x_2^2x_1^2 + \frac{9}{2}\, x_2^4$$

where $k = \int_{-1}^{1} x_1^2 d\xi_1^*(x_1)$. Thus,

$$\max_{x_2} d(x_1,x_2;\ \xi) = 3 \qquad \text{if} \qquad x_1^2 \le k$$

$$= \frac{3}{2}(1 + \frac{x_1^2}{k}) \qquad \text{if} \qquad x_1^2 \ge k$$

and

$$\int_{-1}^{1} \max_{x_2} d(x_1,x_2;\ \xi)d\xi_1^*(x_1) = 3 + \int_{|x_1|<\sqrt{k}} \frac{3}{2}(1-\frac{x_1^2}{k})d\xi_1^*(x_1) > 3 .$$

Therefore, by Theorem 2, $\xi$ is not a marginally restricted D-optimal design.

## 3. Generating Marginally Restricted D-Optimal Designs.

For every $\xi \varepsilon C$ there exists a conditional design $\xi_{2|1}$ such that

$$\xi(x_1,x_2) = \xi_{2|1}(x_2|x_1)\ \xi_1^*(x_1)$$

for all $x_1 \varepsilon S_1^*$ and $\xi_{2|1}$ is an unrestricted design on $x_1 x \chi_2$. The following lemma shows a parallel between the designs $\xi_{2|1}$ and D-optimal designs and indicates how to generate marginally restricted D-optimal designs.

<u>Lemma 4</u>: $\hat{\xi}$ is a marginally restricted D-optimal design if and only if

$$\max_{x_2} d(x_1,x_2;\hat{\xi}) = \int_{\chi_2} d(x_1,x_2;\hat{\xi})\ d\hat{\xi}_{2|1}(x_2|x_1)$$

for all $x_1 \varepsilon S_1^*$.

<u>Proof</u>: For all $\xi \varepsilon C$ and $x_1 \varepsilon S_1^*$ $\max_{x_2} d(x_1,x_2;\xi) \geq \int_{\chi_2} d(x_1,x_2;\xi)\ d\xi_{2|1}(x_2|x_1)$ .

Sufficiency follows by letting $\xi$ in this expression be a marginally restricted D-optimal design, integrating both sides with respect to $\xi_1^*$ and then noting that in the resulting expression the left hand side equals $p$ by (iii) of Theorem 2 and the right hand side equals $p$ by construction.

To show necessity, choose $\tilde{\xi} \varepsilon C$ such that

$$\max_{x_2} d(x_1,x_2;\tilde{\xi}) = \int_{\chi_2} d(x_1,x_2;\tilde{\xi})\ d\tilde{\xi}_{2|1}(x_2|x_1)$$

for all $x_1 \varepsilon S_1^*$ . The result follows by integrating both sides with respect to $\xi_1^*$ and using (iii) of Theorem 2.

Lemma 4 shows that for any $x_1 \varepsilon S_1^*$ we must have

$$d(x_1,x_2;\hat{\xi}) = d(x_1,x_2^*;\hat{\xi})$$

where $x_2$ and $x_2^*$ are points of support of $\hat{\xi}_{2|1}(x_2|x_1)$ . Thus, to iteratively generate marginally restricted D-optimal designs we focus on the conditional designs $\xi_{2|1}$ .

Let $\xi$ denote a design at some iteration. The design $\xi_{+1}$, say, at the next iteration is obtained by augmenting $\xi$ with a fixed point $(x_1^*, x_2^*)$ where $x_1^* \in S_1^*$ and $x_2^* \in \chi_2$ . Specifically, for $0 < \alpha < 1$ let

$$\xi_{2|1}^{\alpha}(x_2|x_1^*) = (1-\alpha)\,\xi_{2|1}(x_2|x_1^*) + \alpha\delta(x_1^*, x_2^*)$$

where $\delta$ places mass 1 at $(x_1^*, x_2^*)$ . The design at the next iteration is now defined as

$$\xi_{+1}(x_1,x_2) = \begin{cases} \xi_{2|1}(x_2|x_1)\ \xi_1^*(x_1) & x_1 \neq x_1^* \\ \xi_{2|1}^{\alpha}(x_2|x_1^*)\ \xi_1^*(x_1^*) & x_1 = x_1^* \end{cases}$$

The following lemma indicates how $(x_1^*, x_2^*)$ is to be chosen.

<u>Lemma 5</u>: Let $\xi$ be any nonsingular marginally restricted design. Define $\xi_{+1}$ as above with $(x_1^*, x_2^*)$ such that

$$\max_{x_1}\,[\max_{x_2} d(x_1, x_2; \xi) - \int_{\chi_2} d(x_1,x_2;\xi)\, d\xi_{2|1}(x_2|x_1)] =$$
$$d(x_1^*, x_2^*; \xi) - \int_{\chi_2} d(x_1^*, x_2; \xi)\, d\xi_{2|1}(x_2|x_1^*) \quad .$$

Then

$$\frac{\partial}{\partial\alpha} \ln|\underline{M}(\xi_{+1})| \Bigg|_{\alpha=0} \geq 0$$

with equality if and only if $\xi$ is a marginally restricted D-optimal design.

Proof: From the definition of $\xi_{+1}$,

$$\underline{M}(\xi_{+1}) = \underline{M}(\xi) + \alpha\, \xi_1^*(x_1^*)\ [\underline{f}(x_1^*,x_2^*)\ \underline{f}'(x_1^*, x_2^*) - \underline{M}(\xi_{2|1}, x_1^*)]$$

where $\underline{M}(\xi_{2|1}, x_1^*) = \int_{\chi_2} \underline{f}(x_1^*, x_2)\ \underline{f}'(x_1^*, x_2)\ d\xi_{2|1}(x_2|x_1^*)$ .

Thus,

$$\frac{\partial}{\partial\alpha} \ln|\underline{M}(\xi_{+1})| = \mathrm{Tr}\ \underline{M}^{-1}(\xi_{+1})\ \ \xi_1^*(x_1^*)\ [\underline{f}(x_1^*, x_2^*)\ \underline{f}'(x_1^*,x_2^*) - \underline{M}(\xi_{2|1}, x_1^*)] \ .$$

(See, for example, Fedorov, 1972).

It follows then that

$$\frac{\partial}{\partial\alpha} \ln|\underline{M}(\xi_{+1})| \Big|_{\alpha=0} = \xi_1^*(x_1^*)\ [d(x_1^*, x_2^*; \xi) - \int_{\chi_2} d(x_1^*, x_2; \xi)\ d\xi_{2|1}(x_2|x_1^*)]$$

Thus,

$$\frac{\partial}{\partial\alpha} \ln|\underline{M}\ (\xi_{+1})| \Big|_{\alpha=0} \geq 0$$

and by Lemma 4 equality is achieved if and only if $\xi$ is a marginally restricted D-optimal design. This completes the proof.

The method of choosing $(x_1^*, x_2^*)$ and generating $\xi_{+1}$ are the essential ingredients in an iterative scheme to generate a marginally restricted D-optimal design. The sequence of weights $\{\alpha_i\}$ and the termination criterion can be specified generally as in schemes for generating unrestricted D-optimal designs. See, Fedorov (1972) and Tsay (1976).

## References

Fedorov, V.V. (1972). The Theory of Optimal Experiments. Translated and edited by W.J. Studden and E.M. Klimko. New York.

Harville, D. (1974). Nearly optimal allocation of experimental units using observed covariate values. Technometrics 16, 589.

Harville, D. (1975). Computing optimum designs for covariance models. In A Survey of Statistical Design and Linear Models. Edited by J.N. Srivastava. North-Holland.

Hoel, P.G. (1965). Minimax designs in two dimensional regression. Ann. Math. Statist. 36, 1097.

Kiefer, J. and Wolfowitz (1960), "The equivalence of two extremum problems," The Canadian Journal of Mathematics, 12, 363-366.

Tsay, J. (1976). On the sequential construction of D-optimal designs. Journal of the American Statistical Association 71, 671-674.