

UCLA

Department of Statistics Papers

Title

Heteroscedastic G-optimal Designs

Permalink

<https://escholarship.org/uc/item/89m2g631>

Author

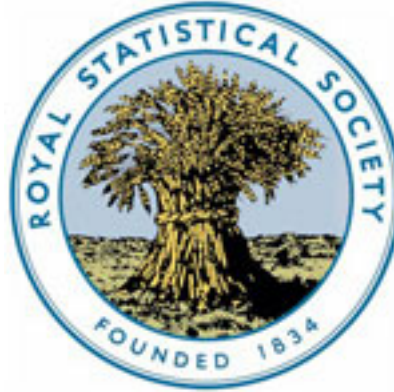
R. Dennis Cook

Publication Date

2011-10-24



**WILEY-
BLACKWELL**



Heteroscedastic G-Optimal Designs

Author(s): Weng Kee Wong and R. Dennis Cook

Source: *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 55, No. 4 (1993), pp. 871-880

Published by: [Blackwell Publishing](#) for the [Royal Statistical Society](#)

Stable URL: <http://www.jstor.org/stable/2345999>

Accessed: 18/05/2011 18:34

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=black>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Blackwell Publishing and Royal Statistical Society are collaborating with JSTOR to digitize, preserve and extend access to Journal of the Royal Statistical Society. Series B (Methodological).

<http://www.jstor.org>

Heteroscedastic G-optimal Designs

By WENG KEE WONG†

and

R. DENNIS COOK

University of California, Los Angeles, USA

University of Minnesota, St Paul, USA

[Received September 1991. Revised June 1992]

SUMMARY

This paper addresses the issue of constructing large sample G-optimal designs when the variability of the response varies across a compact design space. A useful characterization theorem is presented along with a computer algorithm for generating (heteroscedastic) G-optimal designs. To facilitate comparisons between D- and G-optimal designs, Atwood's inequality for comparing D- and G-efficiencies in homoscedastic models is generalized to heteroscedastic models. Some robustness properties of these designs are presented.

Keywords: DESIGN EFFICIENCY; D OPTIMAL; G OPTIMAL; HETEROSCEDASTICITY; INFORMATION MATRIX; LARGE SAMPLE DESIGN

1. INTRODUCTION

There is an abundance of published work on constructing and studying properties of large sample D-optimal designs; see for example Atwood (1969), Kiefer (1985) and Fedorov (1972). One of the common assumptions made in these papers is that the errors are homoscedastic. Consequently, by virtue of the Kiefer and Wolfowitz (1960) theorem, D-optimal designs are also G optimal. When this assumption is not tenable, the equivalence of the two optimal designs does not hold and a different methodology may be required for constructing these designs. For D-optimal designs, a rescaling of the model equation will transform the problem into one in which the errors are homoscedastic, provided that the variance structure of the errors is known. However, there is no such simple gimmick for constructing G-optimal designs when the errors are heteroscedastic. To our knowledge, no explicit treatment of a heteroscedastic G-optimal design has been published.

The intent of this paper is to address the specific issue of constructing these designs and to discuss some of their properties. A more general treatment of constructing minimax optimal designs is given in Wong (1992) but G-optimal designs are particularly useful tools in response surface estimation and we feel that they deserve a more thorough study.

We have p linearly independent continuous regression functions $f^T(x) = (f_1, f_2, \dots, f_p)$ defined on some compact subset Ω of R^k . For each x in Ω , a univariate response variable $y(x)$ is observed in accordance with the statistical model

$$y(x) = f^T(x)\beta + e/\sqrt{\lambda(x)} = \sum_{i=1}^p f_i(x)\beta_i + e/\sqrt{\lambda(x)}. \quad (1.1)$$

†*Address for correspondence:* Department of Biostatistics, School of Public Health, University of California at Los Angeles, 10833 Le Conte Avenue, Los Angeles, CA 90024-1772, USA.

Here, $\beta^T = (\beta_1, \beta_2, \dots, \beta_p)$ are the parameters and the ϵ_s are uncorrelated real-valued error variables having mean 0 and constant variance. It is also assumed that the efficiency function $\lambda(x)$ is a known, bounded, positive real-valued continuous function defined on Ω . Only large sample designs, which are in essence probability measures defined on Ω , are considered. The support of a design ξ is denoted by $\text{supp}(\xi)$ and the set of all large sample designs on Ω is denoted by Ξ .

A common measure of the information contained in a design ξ is the information matrix of the design

$$M(\xi) = \int_{\Omega} f(x) f^T(x) \lambda(x) \xi(dx).$$

Many popular optimality criteria are formulated in terms of the information matrix. In this paper, we are only concerned with designs whose information matrices are non-singular. The variance function of ξ at the point x is defined to be $d(x, \xi) = f^T(x) M(\xi)^{-1} f(x)$ and a G-optimal design is one which minimizes $\bar{d}(\xi) = \max_{x \in \Omega} \{d(x, \xi)\}$.

A design problem is characterized by an ordered triplet $(\Omega, f(x), \lambda(x))$. In our study, it transpires that for a given ξ an important class of designs is the class having support on $A(\xi)$, the set of points in Ω which maximizes $d(x, \xi)$. This subclass of designs is denoted by Ξ_{ξ} . The design which distributes all its mass equally at the extremal points of $\Omega = [-1, 1]^k$ is denoted by η^* and the design which puts all its mass at the point a is denoted by δ_a . Thus, if $k = 1$, η^* can be written as $\frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$.

In the next section, we discuss the methodology for constructing heteroscedastic G-optimal designs. Section 3 gives some general relationships between G- and D-optimal designs in terms of design efficiency, including a generalization of Atwood's (1969) inequality. A concrete example is worked in Section 3. Robustness properties of G-optimal designs with respect to changes in the design space and variation in $\lambda(x)$ are presented in Section 4. We conclude with an algorithm for generating these designs and report the result for a particular design problem when this algorithm is executed. Brief statements are given to indicate how certain results can be justified. In all cases, full details are available in Wong (1990) and Wong and Cook (1990).

2. CONSTRUCTION OF LARGE SAMPLE G-OPTIMAL DESIGNS

The starting point for our study is the minimax characterization theorem given in Fedorov (1980). He considered a rather general abstract problem of finding optimal designs when the optimality criterion is of the minimax type. His result is theoretically interesting but requires further development for application. To re-examine Fedorov's theorem, let $(\Omega, f(x), \lambda(x))$ be a given design problem and let the covariance function of ξ at the point x and u in Ω be denoted by $d(x, u, \xi) = f^T(x) M(\xi)^{-1} f(u)$. We shall also need the squared covariance function $\text{SC}(x, u, \xi): \Omega \times A(\xi) \rightarrow R^1$ defined by $\text{SC}(x, u, \xi) = \lambda(x) d^2(x, u, \xi) - \bar{d}(\xi)$ and the mean-squared covariance (MSC) function $\text{MSC}(x, \mu, \xi): \Omega \times \Xi_{\xi} \rightarrow R^1$ with respect to a probability measure $\mu \in \Xi_{\xi}$ defined by

$$\text{MSC}(x, \mu, \xi) = \int_{A(\xi)} \text{SC}(x, u, \xi) \mu(du).$$

It is straightforward to show that Fedorov's result can be paraphrased as in the following lemma.

Lemma 1. A necessary and sufficient condition for a non-singular design ξ to be G optimal is the existence of a probability measure μ^* on $A(\xi)$ such that

$$MSC(x, \mu^*, \xi) \leq 0 \quad \text{for all } x \text{ in } \Omega. \tag{2.1}$$

Fedorov (1980) did not discuss or offer any hints on how to find an appropriate μ^* when ξ is optimal. However, inequality (2.1) is a useful starting point for deriving an effective strategy for determining μ^* . It is readily checked that

(a) $\max_{y \in \Omega} \{MSC(y, \mu^*, \xi)\} = MSC(x, \mu^*, \xi) = 0 \quad \text{if } x \in \text{supp}(\xi), \tag{2.2}$

(b) when $\lambda(x) = 1$ on Ω , μ^* can be taken to be ξ by virtue of the Kiefer–Wolfowitz theorem and inequality (2.1) reduces to the familiar checking condition for D-optimality (see Fedorov (1972), p. 71).

In general, there may be more than a single μ^* for an optimal ξ . An advantage of formulation (2.1) is that some properties of G-optimal designs can be readily deduced; if ξ is G optimal for the design problem $(\Omega, f(x), \lambda(x))$, then ξ is also G optimal for the design problem $(\Omega, t(x), c \lambda(x))$ where $t(x) = Lf(x)$, L is any $p \times p$ non-singular matrix and $c > 0$.

To determine whether a given design is G optimal, we must verify the existence of a measure μ^* that satisfies the condition in lemma 2.1. There are two methods of doing this. The first method is motivated from equation (2.2) which obviously suggests that we restrict the search for μ^* to only Ξ_ξ instead of arbitrarily selecting an element in Ξ and testing whether inequality (2.1) holds. The second method is a generally more useful technique and is derived again from equation (2.2) by noticing that for any trial design ξ on Ξ and any μ defined on $A(\xi)$ we have

$$\begin{aligned} \max_{x \in \Omega} \left\{ \int_{A(\xi)} \lambda(x) d^2(x, u, \xi) \mu(du) \right\} &\geq \int_{\Omega} \int_{A(\xi)} \lambda(x) d^2(x, u, \xi) \mu(du) \xi(dx) \\ &= \int_{A(\xi)} d(u, \xi) \mu(du) = \bar{d}(\xi). \end{aligned} \tag{2.3}$$

Therefore, μ^* minimizes the left-hand side of the inequality in expression (2.3) when ξ is G optimal and we may characterize μ^* as follows. If ξ is G optimal, then

$$\mu^* = \arg \min_{\mu \in \Xi_\xi} \max_{x \in \Omega} \{MSC(x, \mu, \xi)\} = \arg \min_{\mu \in \Xi_\xi} \max_{x \in \Omega} \{g^T(x) I(\xi, \mu) g(x)\} \tag{2.4}$$

where $g(x) = \lambda(x)^{1/2} M(\xi)^{-1} f(x)$ and $I(\xi, \mu) = \int_{A(\xi)} f(u) f^T(u) \mu(du)$ is the information matrix of a design μ defined on the homoscedastic design space $A(\xi)$.

Let $(\Omega, f(x), \lambda(x))$ be a given design problem. Given a design ξ on Ω , we call any probability measure defined on Ξ_ξ an A -measure associated with ξ if

$$\mu_{A, \xi} = \arg \min_{\gamma \in \Xi_\xi} \max_{x \in \Omega} \{MSC(x, \gamma, \xi)\}.$$

The above argument shows that the problem of finding a G-optimal design is closely connected to that of finding an A -measure associated with the G-optimal design. The latter problem is essentially a linear minimax problem, as can be seen from equation (2.4) and thus should be amenable to simple solutions under appropriate assumptions. For example, it is straightforward to show that if ξ is G optimal for $(\Omega, f(x), \lambda(x))$, $A(\xi) \supseteq \text{supp}(\xi)$ and $\text{supp}(\xi) = [y | \lambda(y) = \max_{x \in \Omega} \{\lambda(x)\}]$, then $\mu_{A, \xi} = \xi$.

Another simple case where $\mu_{A,\xi}$ can be determined analytically is as follows. Suppose that $\Omega = [-1, 1], f^T(x) = (1, x)$ and $\lambda(x)$ is symmetric. To see that $\mu_{A,\xi} = \eta^*$, let $x^* = \arg \max\{\lambda(x) (a^2 + b^2x^2)\}$ and let $\mu_{A,\xi} = p\delta_1 + (1-p)\delta_{-1}$ for some $0 \leq p \leq 1$. (This is possible because $d(x, \xi)$ is a quadratic function of x and thus $A(\xi)$ is a subset of $[-1, 1]$ for any non-singular design ξ .) The conclusion follows from the symmetry of $\lambda(x)$ and the fact that either

$$2ab(1 - 2p)|x^*| \geq 0 \tag{2.5a}$$

or

$$-2ab(1 - 2p)|x^*| \geq 0 \tag{2.5b}$$

for any positive numbers a, b and any p , with equality in both inequalities when $p = 0.5$.

As will be shown in Section 3, inequality (2.1) together with the identification of $\mu_{A,\xi}$ allows us a simple way to verify whether a design ξ is G optimal. The $MSC(x, \mu_{A,\xi}, \xi)$ function for a candidate design ξ is plotted and its proximity to the optimal design is assessed from the graph.

3. EFFICIENCY OF G-OPTIMAL DESIGNS

In this section, it is helpful to distinguish whether we are assuming that $\lambda(x)$ is constant or not; we do this by deliberately inserting λ into our usual notation. For example, the variance function is now denoted by $d(x, \xi, \lambda)$ instead of $d(x, \xi)$ to emphasize that we are now working under a heteroscedastic model with efficiency function $\lambda(x)$. Thus, $d(x, \xi, 1)$ and $M(\xi, 1)$ are respectively the variance function and the information matrix of ξ in a homoscedastic model. Accordingly, we also let ξ_G^λ and ξ_D^λ denote the G- and D-optimal designs for the design problem $(\Omega, f(x), \lambda(x))$. It is intuitive to expect ξ_G^λ and ξ_D^λ to be close when $\lambda(x)$ is nearly constant by the Kiefer-Wolfowitz theorem. To formalize this idea, recall that the D-efficiency of a design η relative to a non-singular design ξ is

$$D(\eta, \xi, \lambda) = \{|M(\eta, \lambda)| / |M(\xi, \lambda)|\}^{1/p}.$$

When $\xi = \xi_D^\lambda$, we use the abbreviated notation $\text{eff}_D(\eta, \lambda) = D(\eta, \xi_D^\lambda, \lambda)$ and call this the D-efficiency of the design η . The G-efficiency of a non-singular design η relative to a non-singular design ξ is

$$G(\eta, \xi, \lambda) = \bar{d}(\xi, \lambda) / \bar{d}(\eta, \lambda).$$

When $\xi = \xi_G^\lambda$, we set $\text{eff}_G(\eta, \lambda) = G(\eta, \xi_G^\lambda, \lambda)$ and call this the G-efficiency of the design η . For homoscedastic models, $\text{eff}_G(\eta, 1) = p / \bar{d}(\eta, 1)$ by the Kiefer-Wolfowitz theorem.

It is certainly useful to be able to quantify how close a given design is to the optimal design, without knowing the latter. This could be done graphically by using the MSC function with the aid of the following result.

Theorem 1. Let $(\Omega, f(x), \lambda(x))$ be a given design problem. Suppose that ξ is a non-singular design, μ is its associated A -measure and

$$\max_{x \in \Omega} \{MSC(x, \mu, \xi)\} = \epsilon \quad \text{for some } \epsilon \geq 0. \tag{3.1}$$

Then,

$$\text{eff}_G(\xi, \lambda) \geq 1 - \epsilon/\bar{d}(\xi, \lambda). \tag{3.2}$$

The proof of this result relies on the convexity of the criterion. For most practical purposes, the experimenter is willing to use a design if the ϵ in equation (3.1) is sufficiently small. Evidently, the case $\epsilon = 0$ corresponds to the case when ξ is optimal. Theorem 1 thus establishes the relationship between ϵ and the efficiency of the design.

Although the above method gives us a practical way of verifying whether a design is G optimal, it may be reasonable in practice to use a D-optimal design as an approximation. The reason for this is that generating D-optimal designs is relatively easy and they have several useful optimal properties. Denoting the maximum and minimum values of $\lambda(x)$ on Ω by λ_{\max} and λ_{\min} respectively, we have the following theorem.

Theorem 2.

$$\begin{aligned} \text{(a)} \quad & \text{eff}_G(\xi_D^\lambda, \lambda) \geq \lambda_{\min}/\lambda_{\max}. \\ \text{(b)} \quad & \text{eff}_D(\xi, \lambda) \geq (\lambda_{\min}/\lambda_{\max}) \text{eff}_G(\xi_G^\lambda, 1) \text{eff}_G(\xi, \lambda). \end{aligned} \tag{3.3}$$

Note that when $\lambda(x) = 1$ inequality (3.3) reduces to Atwood's (1969) inequality. In addition, when $\xi = \xi_G^\lambda$, we have

$$\text{eff}_D(\xi_G^\lambda, \lambda) \geq (\lambda_{\min}/\lambda_{\max}) \text{eff}_G(\xi_G^\lambda, 1),$$

which is the D-efficiency counterpart to theorem 2, part (a).

Theorem 2, part (a), gives a lower bound for the G-efficiency of using ξ_D^λ instead of ξ_G^λ . As expected, when $\lambda(x)$ is constant on Ω , $\text{eff}_G(\xi_D^\lambda, \lambda) = 1$ and D- and G-optimal designs are equivalent. The practical implication of theorem 2, part (b), is that, under the assumption of heteroscedasticity, Atwood's inequality becomes invalid and it is no longer true that finding a design to maximize its G-efficiency necessarily provides a lower bound for its D-efficiency. In this case, the experimenter may have to decide carefully which optimality criterion is more meaningful.

As an example, consider finding the G-optimal design when $\Omega = [-1, 1]$, $f^T(x) = (1, x)$ and $\lambda(x) = 1.5 - |x|$. Using Kiefer-Wolfowitz's theorem, it can be verified that the D-optimal design for this problem is $\xi_D = 0.5\delta_{0.75} + 0.5\delta_{-0.75}$ and $\bar{d}(\xi_D) = 3.7037$. By lemma 1 and the argument used in the paragraph containing inequalities (2.5), we need only to examine the behaviour of the function $\text{MSC}(x, \eta^*, \xi_D)$ to verify whether ξ_D is also G optimal. From the plot of the function

$$\text{MSC}(x, \eta^*, \xi_D) = 1.7778(1.5 - |x|)(1 + 3.1605x^2) - 3.7037$$

in Fig. 1, it is evident that ξ_D is not G optimal since it has a positive maximum value of 0.0962 at ± 0.8802 . Nevertheless, ξ_D has a G-efficiency of at least $1 - 0.0962/3.7037 = 0.9740$ by theorem 1. The G-optimal design for this problem turns out to be $\xi_G = 0.5\delta_a + 0.5\delta_{-a}$, where a solves the equation $a^3 + 3a - 3 = 0$ and is approximately equal to 0.8177. The optimality of ξ_G can be verified by noting that the function

$$\text{MSC}(x, \eta^*, \xi_G) = 2.1481(1.5 - |x|)(1 + 2.2368x^2) - 3.6576$$

is non-positive for all $x \in [-1, 1]$ and is equal to 0 at ± 0.8177 (see Fig. 1). Note that $\bar{d}(\xi_D) - \bar{d}(\xi_G) = 0.0461$ so that the high G-efficiency of ξ_D is not really surprising. See Wong (1990) for more complicated examples.

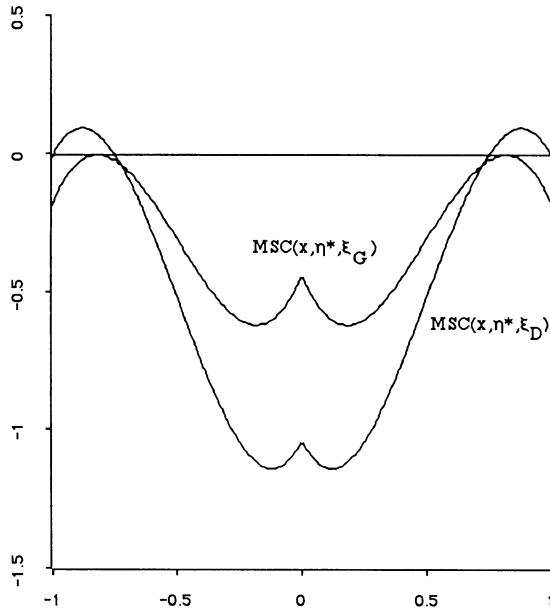


Fig. 1. Plots for $MSC(x, \eta^*, \xi_D)$ and $MSC(x, \eta^*, \xi_G)$

4. ROBUSTNESS PROPERTIES

In many practical situations, it may be problematic to specify $\lambda(x)$. A wrong specification of $\lambda(x)$ may result in a different optimal design. In this section, we consider how the G-optimal design varies when $\lambda(x)$ is changed in a specified way and when Ω is transformed linearly. The second case is useful because it will help us to deduce optimal designs in different design spaces. For this section, we shall assume that Ω is symmetric: $x \in \Omega$ implies $-x \in \Omega$. We also need to define two new constructions:

- (a) given a design $\xi = \sum_{i=1}^s \xi_i \delta_{x_i}$, we define $\bar{\xi} = \sum_{i=1}^s \xi_i \delta_{-x_i}$ and call the latter design the mirror image of ξ ;
- (b) corresponding to an efficiency function $\lambda(x)$ on Ω , we define $\bar{\lambda}(x) = \lambda(-x)$ —obviously, $\bar{\lambda}(x)$ is also an efficiency function and is equal to $\lambda(x)$ if $\lambda(x)$ is an even function.

Theorem 3. Let ξ be G optimal for the design problem $(\Omega, f(x), \lambda(x))$, where Ω is symmetric. Suppose that there exists a matrix $L \in \{\text{diag}(d_1, d_2, \dots, d_p) \mid d_i = \pm 1, i = 1, 2, \dots, p\}$ such that $f(x) = Lf(-x)$ for all $x \in \Omega$. Then, the mirror image of ξ is G optimal for the design problem $(\Omega, f(x), \bar{\lambda}(x))$.

The justification of theorem 3 hinges on the three conditions

- (a) $M(\bar{\xi}) = LM(\xi)L^T$,
- (b) $d(x, \xi) = d(-x, \xi)$ for all $x \in \Omega$ and
- (c) $d(x, -u, \xi) = d(-x, u, \xi)$ for $u \in A(\xi)$.

The hypothesis imposed in theorem 3 is not restrictive. It includes the class of all polynomial and trigonometric regression functions, as well as improper polynomial regression functions like $f^T(x) = (1, x, x^3)$ etc. For simple linear regression models with $f^T(x) = (1, x)$, a stronger result is possible: if $\lambda(x)$ satisfies

$$\lambda(x) \left\{ \frac{(1+x)^2}{\lambda(1)} + \frac{(1-x)^2}{\lambda(-1)} \right\} \leq 4 \quad \text{for all } x \in \Omega,$$

then

$$\xi = \frac{\lambda(-1)}{\lambda(-1) + \lambda(1)} \delta_1 + \frac{\lambda(1)}{\lambda(-1) + \lambda(1)} \delta_{-1}$$

is G optimal for the problem $(\Omega, f(x), \lambda(x))$ and $\bar{\xi}$ is G optimal for the problem $(\Omega, f(x), \bar{\lambda}(x))$.

We say that an optimal design is scale invariant to changes in $\lambda(x)$ if it is optimal for both the problems $(\Omega, f(x), \lambda(x))$ and $(\Omega, f(x), c \lambda(x))$, $c > 0$. In the same spirit, we say that an optimal design is location invariant to changes in $\lambda(x)$ if it remains optimal for all $c \geq 0$ in the problem $(\Omega, f(x), \lambda(x) + c)$. By an earlier remark (see the paragraph after equation (2.2)), G-optimal designs and D-optimal designs have this scale invariant property. However, neither D- nor G-optimal designs are location invariant to changes in $\lambda(x)$. For example, if $\Omega = [-1, 1]$ and $f^T(x) = (1, x)$, then η^* is both D and G optimal for $\lambda(x) = c - |x|$ if and only if $c \geq 2$. However, our experience with heteroscedastic D- and G-optimal designs suggests that the location effect is important only when there are regions in the design space when $\lambda(x)$ is near to 0. Specifically, there seems to be a critical number λ_c associated with each design problem such that, if $\lambda_{\min} \geq \lambda_c$, then the optimal design for the problem $(\Omega, f(x), \lambda(x))$ is the same for all other problems of the form $(\Omega, f(x), \lambda(x) + c)$ where $c \geq \lambda_c$. The implication of this is that, if we know the shape of the efficiency function and the value of λ_c , then the heteroscedastic errors can be treated as if they were homoscedastic as far as constructing D- or G-optimal designs is concerned.

The next result concerns the behaviour of the G-optimal design when Ω is subjected to a linear transformation.

Theorem 4. Let ξ be G optimal for the design problem $(\Omega, f(x), \lambda(x))$ and τ be a class of one-to-one transformations from Ω to $\bar{\Omega}$. Assume that for each $g \in \tau$ there is a non-singular $p \times p$ matrix \bar{g} such that

$$f(gx) = \bar{g}f(x) \quad \text{for all } x \in \Omega \text{ and } g \in \tau. \tag{4.1}$$

Define $\bar{\lambda}(y) = \lambda(g^{-1}y) = \lambda(x)$ on $\bar{\Omega}$ and associate each design ξ on Ω with a corresponding design $g\xi$ on $\bar{\Omega}$ defined by

$$g\xi(A) = \xi(g^{-1}A) \quad \text{for all Borel subsets } A \text{ of } \bar{\Omega}.$$

Then the G-optimal design for the problem $(\bar{\Omega}, f(y), \bar{\lambda}(y))$ is $g\xi$.

The key idea to the proof of theorem 4 lies in noting that under assumption (4.1)

- (a) $M(g\xi) = \bar{g}M(\xi)\bar{g}^T$ and
- (b) $d(gx, g\xi) = d(x, \xi)$ for all $x \in \Omega$ and $g \in \tau$.

The hypothesis imposed on $f(x)$ is restrictive and should be kept in mind when we ‘move’ designs from one design space to another; it appears that this observation,

which is relevant even if $\lambda(x)=1$, has been overlooked. To see the implication of theorem 4, observe that, if τ is the set of all $k \times k$ non-singular matrices, then any proper polynomial regression functions of the form $(1, x, x^2, \dots, x^{p-1})$ can be shown to satisfy equation (4.1). However, if $\tau = \tau_1 = \{Ax + b \mid A \text{ is non-singular and } b \neq 0\}$, then equation (4.1) is violated for any improper polynomial regression function, i.e. polynomial regression functions with missing terms. For example, if $\Omega = [-1, 1]$, $\lambda(x)=1$ and $f(x)=x$, δ_{-1} is D (or G) optimal. (Indeed, any design of the form $p\delta_{-1} + (1-p)\delta_1$ ($0 \leq p \leq 1$) is D (or G) optimal for this problem.) Consider $g \in \tau_1$ where $gx = (x+1)/2$. Then $g\delta_{-1} = \delta_0$ but clearly the design δ_0 is not D (or G) optimal for the design problem $([0, 1], f(y)=y, 1)$. In short, the practice of designing on a standardized design space and then transforming the design for the problem at hand is valid only under restrictive conditions.

5. COMPUTER ALGORITHM

To apply this design theory to a wide variety of practical problems, it is essential to have a computer algorithm to search for the optimal design since optimal designs derived solely from theoretical considerations are usually limited in scope and applicability. Some basic computer algorithms for generating various kinds of optimal design are described in Mitchell (1974) and Fedorov (1972).

We propose the following algorithm for generating heteroscedastic G-optimal designs. It is based largely on heuristics and has several common features with the current algorithms for finding D- and L-optimal designs. It is similar also to the algorithm in Wong (1990) for generating E-optimal designs. The convergence of this algorithm to the optimal design is not addressed because theorem 1 can be used to check the proximity of the candidate design from the optimum. Since this algorithm is iterative, we shall use v_k to denote the number of elements in $A(\xi_k)$ at the k th iteration.

- (a) Step 0: set $k=0$ and generate a non-singular starting design ξ_k . Let x^* be an arbitrary point in $\text{supp}(\xi_k)$.
- (b) Step 1: find the elements u_1, u_2, \dots, u_{v_k} in $A(\xi_k)$.
- (c) Step 2: determine an A -measure $\mu_k = \sum_{i=1}^{v_k} \mu_i \delta_{u_i}$ for ξ_k by solving the following linear programming problem. Minimize

$$z = \sum_{i=1}^{v_k} \mu_i \lambda(x^*) \{f^T(x^*) m^{-1}(\xi_k) f(u_i)\}^2$$

with respect to $\mu_1, \mu_2, \dots, \mu_{v_k}$ subject to $\sum_{i=1}^{v_k} \mu_i = 1$ and $\mu_i \geq 0, i=1, 2, \dots, v_k$.

- (d) Step 3: check whether $\max_{x \in \Omega} \{MSC(x, \mu_k, \xi_k)\} \leq \epsilon$ where ϵ is a preassigned small positive number of order say 0.001. If yes, stop and conclude that ξ_k is G optimal for the given tolerance level ϵ . Otherwise, go to step 4.
- (e) Step 4: a new point x_k^* is added to the current design ξ_k . Following standard optimal design strategies, x_k^* is chosen so that

$$MSC(x_k^*, \mu, \xi_k) = \max_{x \in \Omega} \{MSC(x, \mu_k, \xi_k)\}$$

and μ_k is an A -measure for ξ_k determined in step 2. This gives a new design

TABLE 1

G-efficiencies of designs generated by the algorithm for $\Omega = [-1, 1]$, $f^T(x) = (1, x)$ and $\lambda(x) = \exp(-x/2)$ with starting design η^* and three weighting sequences †

<i>k</i>	Efficiencies for the following weighting sequences:					
	$\alpha_k = 1/(k+1)$	$\alpha_k = 1/\sqrt{k+1}$	$\alpha_k = 1/\ln(k+1)$	$\alpha_k = 1/(k+1)$	$\alpha_k = 1/\sqrt{k+1}$	$\alpha_k = 1/\ln(k+1)$
10	0.929570	0.945543	0.927529	0.792028	0.999765	0.975302
50	0.998552	0.990048	0.864730	0.928850	0.995388	0.997386
100	0.991713	0.994941	0.999137	0.946602	0.998861	0.998890
200	0.995132	0.997486	0.947403	0.965581	0.999151	0.999523
500	0.999920	0.998981	0.999279	0.977692	0.999908	0.999832

†The first of the two columns is the efficiency attained at the *k*th iteration and the second column is the arithmetic mean of $\text{eff}_D(\xi_j, \lambda)$, $j = k-5, k-4, \dots, k+5$.

$\xi_{k+1} = (1 - \alpha_k)\xi_k + \alpha_k \delta_{x_k^*}$, where α_k is some suitable weighting sequence; see Tsay (1976).

(f) Step 5: set *k* to *k* + 1 and go to step 1 with the new design ξ_{k+1} and $x^* = x_k^*$.

We have implemented this algorithm with S (Becker and Chambers, 1984) on a VAX 11/750 computer and tested its performance quite extensively in problems with $\Omega = [-1, 1]$, $f^T(x) = (1, x)$ and different types of efficiency function. The results have always been promising. For all the problems that we have looked at, the algorithm generates a design with very high *G*-efficiency. Table 1 is a typical kind of result when the algorithm is executed. Occasionally, with judicious choice of the α_k , the algorithm produces the *G*-optimal design after only a few iterations. For example, if $\lambda(x) = 1/(x+c)$ or $\lambda(x) = x+c$, $c > 1$ and $\alpha_k = 1/(c+k)$, it can be shown analytically that the algorithm yields the *G*-optimal designs after one iteration.

This proposed algorithm is encouraging and has potential for other applications. However, at present, the mathematical and numerical properties of the algorithm are still not well understood. For example, in step 1, the numerical computation of the elements in $\mathcal{A}(\xi_k)$ may have the same problem encountered in Atkinson and Fedorov (1975). Currently, further study of the algorithm is under way, including applying this algorithm to more complicated design problems. A referee has also suggested that we allow α_k to take on negative values and possibly introduce tolerance intervals into the algorithm, as in Atkinson and Fedorov (1975). This and other related issues will be dealt with in a subsequent paper.

ACKNOWLEDGEMENTS

Much of this paper is taken from Wong's doctoral thesis, written under the supervision of Cook at the University of Minnesota. This work was partially supported by the National Science Foundation grant 90-001298.

REFERENCES

- Atkinson, A. C. and Fedorov, V. V. (1975) Optimal design: experiments for discriminating between several models. *Biometrika*, **62**, 289-304.
 Atwood, C. L. (1969) Optimal and efficient designs for experiments. *Ann. Math. Statist.*, **40**, 1570-1602.

- Becker, R. A. and Chambers, J. M. (1984) *S, an Interactive Environment for Data Analysis and Graphics*. Belmont: Wadsworth.
- Fedorov, V. V. (1972) *Theory of Optimal Experiments* (eds W. J. Studden and E. M. Klimko). New York: Academic Press.
- (1980) Convex design theory. *Math. Operforsch. Statist.*, **11**, 403–413.
- Kiefer, J. (1985) *Jack Carl Kiefer, Collected Papers III: Design of Experiments*. New York: Springer.
- Kiefer, J. and Wolfowitz, J. (1960) The equivalence of two extremum problems. *Can. J. Math.*, **12**, 363–366.
- Mitchell, T. J. (1974) An algorithm for the construction of D-optimal experimental designs. *Technometrics*, **16**, 203–211.
- Tsay, J. Y. (1976) On the sequential construction of D-optimal designs. *J. Am. Statist. Ass.*, **71**, 671–674.
- Wong, W. K. (1990) Heteroscedastic G-optimal designs. *PhD Thesis*. School of Statistics, University of Minnesota, Minneapolis.
- (1992) A unified approach to mini-max optimal designs. *Biometrika*, **79**, 611–619.
- Wong, W. K. and Cook, R. D. (1990) Heteroscedastic G-optimal designs. *Statistics Series 72*. University of California, Los Angeles.