Title
A Unified Approach to Mini-Max Designs Construction

Permalink
https://escholarship.org/uc/item/4pb3x0jc

Author
Weng-Kee Wong

Publication Date
2011-10-24
A unified approach to the construction of minimax designs

BY WENG-KEE WONG

Department of Biostatistics, University of California at Los Angeles,
Los Angeles, California 90024, U.S.A.

SUMMARY

We present a unified approach for constructing large-sample optimal designs when the optimality criterion is of the minimax type. The assumed model is a general linear regression model with a known efficiency function defined on a closed and bounded space. An equivalence theorem is formulated and computer algorithms for generating minimax optimal designs are proposed. It is shown that this methodology is simple and rather general. It can be used to construct, for example, minimax variance optimal designs, minimax with respect to the single parameters designs and \(E\)-optimal designs. An application of this procedure to find an optimal design for minimizing the maximum predictive variance over a compact region in a heteroscedastic design problem is included.

Some key words: Efficiency function; Heteroscedasticity; Information matrix; Large-sample designs.

1. Introduction

The purpose of this paper is to extend the results of Wong & Cook (1992) to a wide class of design problems and to demonstrate a methodology for finding minimax designs. It appears that while this methodology has been used occasionally in the literature (Atkinson & Fedorov, 1975a, b; Fedorov & Khabarov, 1986), its general applicability and importance have been overlooked. Fedorov (1980) briefly described the construction of minimax optimal designs from a theoretical viewpoint using this approach but the consideration given there is superficial; he did not have a practical way of verifying if a given design is optimal.

Wong & Cook (1992) considered the construction of \(G\)-optimal designs in the presence of heteroscedasticity. In this paper, we show that \(G\)-optimal designs, like many other kinds of minimax optimal designs, can be formulated using this general approach. As will be shown by examples, this methodology is illuminating and provides a straightforward way of verifying and understanding properties of optimal designs. Because of the general nature of the methodology, it can suggest new ideas to answer some outstanding questions. For example, there appears to be no known algorithm for generating \(E\)-optimal designs; our approach will suggest one in §5.

We assume the general linear regression model. Specifically, let \(f_1, f_2, \ldots, f_p\) be \(p\) given linearly independent continuous regression functions defined on some compact design subspace \(\Omega\) of \(R^k\) and let \(f^T(x) = (f_1(x), f_2(x), \ldots, f_p(x))\). For each \(x\) in \(\Omega\), a univariate response variable \(y(x)\) is observed under the model

\[ y(x) = f^T(x)\beta + e/\{\lambda(x)\}^1, \]

where \(\beta^T = (\beta_1, \beta_2, \ldots, \beta_p)\) are the parameters and the \(e\)'s are uncorrelated real-valued error random variables having mean zero and constant variance. It is also assumed that
\( \lambda(x) \) is a known, bounded, positive real-valued continuous function defined on \( \Omega \). The function \( \lambda(x) \) is commonly called the efficiency function and it reflects the heteroscedasticity structure in the model. When \( \lambda(x) \) is constant on \( \Omega \), we may assume without loss of generality that \( \lambda(x) = 1 \) and then we have a homoscedastic model.

Let \( \Xi \) be the set of all probability measures \( \xi \) defined on \( \Omega \). An element of \( \Xi \) is called a large-sample experimental design or, simply, a design. Thus, if the design \( \xi \) has mass \( \xi_i \) at \( x_i \) and \( n \) observations are allowed for the experiment, approximately \( n\xi_i \) observations are taken at \( x_i \). The \( x_i \)'s are called the support points of \( \xi \). If a design has only one support point at \( x_i \) it will be denoted by \( \delta_i \). A common measure of the information contained in a design \( \xi \) is the information matrix of the design defined by

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

We are primarily concerned with nonsingular designs, designs whose information matrices are nonsingular. The unknown parameters \( B \) are estimated by the least squares method and the variance of the fitted value at the point \( x \) using design \( \xi \) is, apart from an unimportant factor,

\[
d(x, \xi) = f^T(x)M^{-1}(\xi)f(x) = \text{tr}\{M^{-1}(\xi)f(x)f^T(x)\}.
\]

A design problem is characterized by the triplet \((\Omega, f(x), \lambda(x))\) together with a convex optimality criterion function \( \Phi \), which is selected to reflect the experimenter's interest. The problem confronting the experimenter is how to select a design so that \( \Phi \) is minimized. Designs which minimize \( \Phi \) are called \( \Phi \)-optimal. Many of the popular optimality criteria \( \Phi \) are formulated in terms of the information matrix; see Kiefer (1974) or Atkinson (1988) for a more comprehensive discussion of this subject.

In this paper, we focus on a special class of optimality criteria. Section 2 discusses this class and gives an equivalence theorem for verifying the optimality of such a design in practice. Section 3 demonstrates in some detail how this approach may be applied to obtain familiar results for \( E \)-optimal designs. In §4, further applications are indicated. These include designs which minimize the maximum variance across a specified region, minimax with respect to single parameters designs and a new type of design called \( P \)-optimal designs useful for prediction purposes. In §5, an algorithm for generating the class of minimax designs considered here is proposed and an example of a \( P \)-optimal design is given.

2. A general equivalence theorem

Given a design problem \((\Omega, f(x), \lambda(x))\), let \( \Phi \) have the form

\[
\Phi(M(\xi)) = \max_{z \in Z} \Psi(z, M(\xi)) \tag{2.1}
\]

for some functional \( \Psi(\cdot) \) and \( Z \) an arbitrary Euclidean compact set. As will be seen, many commonly used optimality criteria are included in (2.1) by different choices of \( Z \). Additionally, we assume, for each \( z \in Z \),

(a) \( \Psi(z, A + B) \leq \Psi(z, A) \) for all nonnegative definite matrices \( A \) and \( B \);

(b) \( \Psi(z, nM(\xi)) = \rho(n)\Psi(z, M(\xi)) \) for some decreasing function \( \rho(\cdot) \);

(c) \( \{M(\xi) | \xi \in \Xi\} \) is a compact set;

(d) \( \Psi(z, M(\xi)) \leq (1 - \alpha)\Psi(z, M(\xi_1)) + \alpha\Psi(z, M(\xi_2)) \), where \( \xi = (1 - \alpha)\xi_1 + \alpha\xi_2 \);
Unified approach to construction of minimax designs

(e) $\Psi(z, M(\xi)) < \infty$ for some $\xi \in \Xi$;
(f) $\Psi(z, M)$ is differentiable on any set of the form

$$\{M(\xi): \Psi(z, M(\xi)) \leq c < \infty, \xi \in \Xi\}.$$  

Under these and some additional regularity conditions, Fedorov (1980) stated without proof an equivalence theorem for a $\Phi$-optimal design $\xi^*$.

**Theorem 2.1.** Let $\lambda(x)$ and $f(x)$ be continuous functions defined on the compact spaces $\Omega$ and $Z$ and assume $\Psi$ satisfies all the conditions (a)–(f). Let $\nabla \Psi(z, M(\xi))$ be the $p \times p$ matrix with elements given by

$$\{\nabla \Psi(z, A)\}_{ij} = \frac{d\Psi(z, A)}{d\{A\}_{ij}}, \quad (i, j = 1, 2, \ldots, p),$$

$$\bar{\Psi}(M(\xi)) = \max_{z \in Z} \Psi(z, M(\xi)), \quad A(\xi) = \{a \in Z | \Psi(a, \xi) = \bar{\Psi}(M(\xi))\}.$$  

Then, a necessary and sufficient condition that $\xi^*$ is $\Phi$-optimal is that there exists a probability measure $\mu^*$ on $A(\xi^*)$ such that, for all $x \in \Omega$,

$$\text{tr} \left\{ \lambda(x) f(x) f^T(x) \int_{A(\xi^*)} \nabla \Psi(u, M(\xi^*)) \mu^*(du) \right\} \geq \text{tr} \left\{ M(\xi^*) \int_{A(\xi^*)} \nabla \Psi(u, M(\xi^*)) \mu^*(du) \right\}. \quad (2.2)$$

A remarkable feature of the proof of Lemma 3.1 of my Minnesota Ph.D. dissertation, which is a special case of Theorem 2.1, is the lack of dependence on the specific form of the $G$-optimality criterion. More generally, the proof can be modified in a straightforward manner to include design problems where $\Psi$ has the form

$$\Psi(z, M(\xi)) = \text{tr} \{B(z) M(\xi)^{-1}\} \quad (2.3)$$

for some nonnegative definite matrix $B(z)$. The equivalence theorem for the optimality criterion function in (2.3) becomes: $\xi^*$ is optimal if and only if there exists a probability measure on $A(\xi^*)$ such that, for all $x \in \Omega$,

$$\text{tr} \left\{ \lambda(x) f(x) f^T(x) \int_{A(\xi^*)} M(\xi^*)^{-1} B(u) M(\xi^*)^{-1} \mu^*(du) \right\} - \text{tr} \{B(a) M(\xi^*)^{-1}\} \leq 0. \quad (2.4)$$

The set $A(\xi)$ is sometimes called the ‘answering set’: the elements $u$ in $A(\xi)$ ‘answer $\xi$’ (Danskin, 1967, p. 21). Here and throughout we will denote a generic element in the answering set by ‘$a$’. In the derivation of (2.4) we have used the fact that

$$\nabla \Psi(u, M(\xi)) = -M^{-1}(\xi) B(u) M^{-1}(\xi);$$

thus, the right-hand side in (2.2) simplifies to

$$-\int_{A(\xi^*)} \text{tr} \{B(u) M^{-1}(\xi^*) \mu^*(du)\} = -\text{tr} \{B(a) M^{-1}(\xi^*)\},$$

which by definition is a constant for any $a \in A(\xi^*)$. The left-hand side of (2.4) is a function involving $x$ only and in practice we would examine its graph to help us verify if a design is optimal.
It is evident from (2-4) that the optimality of a design $\xi^*$ is intimately related to the existence of a certain probability measure $\mu^*$ on $A(\xi^*)$. This suggests that the crux of verifying the optimality of a design lies in the ease of finding $\mu^*$. An advantage in the formulation (2-4) is that we have basically reduced the problem of verifying the optimality of a design to that of finding a probability measure defined on $A(\xi^*)$. The latter problem is essentially a linear one and therefore should be easier to work with from the computational standpoint. However, except for a few design problems which we describe in the next paragraph, a general analytic solution for $\mu^*$ is usually problematic and one needs to resort to an iterative scheme for determining $\mu^*$.

Several commonly used optimality criteria are covered under (2-3). For example, if we are interested in $A$-optimality, we set $Z = \Omega$ and $B(z)$ equal to the $p \times p$ identity matrix. In this case, any probability measure supported on the answering set for the optimal design can be used as $\mu^*$ in (2-4). For $G$-optimality, we have $Z = \Omega$ and $B(z) = f(z)f^T(z)$. Interestingly, for homoscedastic models, there is a simple relationship between the $G$-optimal design $\xi^*$ and $\mu^*$, namely $\mu^* = \xi^*$; see my Ph.D. dissertation. As a further example, if we specify $Z$ to be a singleton set, say $Z = \{z\}$, and let $d(x, u, \xi) = f^T(x)M^{-1}(\xi)f(u)$, the equivalence statement in (2-4) simplifies to

$$\lambda(x)d^2(x, z, \xi^*) - d(z, \xi^*) \leq 0$$

(2-5)

for all $x \in \Omega$, since $A(\xi) = \{z\}$ for any $\xi \in \Xi$ and $\mu^* = \delta_z$. We note that (2-5) is the familiar checking condition for a $L$-optimal design corresponding to the criterion $\text{tr} \left\{BM(\xi)\right\}^{-1}$, with $B = f(z)f^T(z)$. The design $\xi^*$ in (2-5) has the smallest variance at the point $z$ among all other designs in $\Xi$. For homoscedastic polynomial regression with $f^T(x) = (1, x, x^2, \ldots, x^{p-1})$ on $\Omega = [-1, 1]$ and $z \notin \Omega$, this is the usual extrapolation design problem and the $\xi^*$ in (2-5) can be described analytically (Fedorov, 1972, p. 149) by $\xi^* = \Sigma \xi_i \delta_{x_i}$, where $x_1, x_2, \ldots, x_p$ are the $p$ roots of the polynomial $\cos \{\frac{p-1}{2} \cos^{-1} x\}$. $\xi_i$ is given by

$$\xi_i = |l_i(z)|/\sum |l_j(z)| \quad (i = 1, 2, \ldots, p),$$

(2-6)

and the $l_i(x)$'s are the Lagrange interpolation polynomials for the points $x_1, x_2, \ldots, x_p$. We denote this design by $\xi_{L,p}$ and will refer to it again in § 4.

If $\xi^*$ is $\Phi$-optimal, we may limit the search for $\mu^*$ to a subclass of probability measures defined on $A(\xi^*)$. For a given design $\xi$ on $\Omega$, let

$$\mu_{A,\xi} = \arg \min_{\xi \in \Xi} \max_{x \in \Omega} \text{tr} \left\{\lambda(x)f(x)f^T(x) \int_{A(\xi)} M(\xi)^{-1}B(u)M(\xi)^{-1}\xi(du)\right\}$$

(2-7)

and call any probability measure $\mu_{A,\xi}$ that satisfies (2-7) an $A$-measure for the design $\xi$. Here, $\Xi$ is the set of all designs with support on $A(\xi)$. Using a similar argument to that in my dissertation, the optimality of $\xi$ can now be verified by simply integrating

$$\text{tr} \left\{\lambda(x)f(x)f^T(x)M(\xi)^{-1}B(u)M(\xi)^{-1}\right\} - \text{tr} \left\{B(\xi)M^{-1}(\xi)\right\}$$

with respect to any $A$-measure for $\xi$ and observing if (2-4) is satisfied.

In §§ 3 and 4, we apply this technique to some design problems. For expository purposes, all our examples assume that $\Omega = [-1, 1], f^T(x) = (1, x, x^2, \ldots, x^{p-1}), \lambda(x) = 1$ and $Z$ is some compact Euclidean subspace. The checking criterion (2-4) is a general one and, as long as the assumptions laid out in § 1 are satisfied, it can be used to verify the optimality of a design for a broad class of problems, including heteroscedastic models and situations where $Z$ is some union of compact sets. The methods employed by Ehrenfeld (1955), Gaylor & Sweeney (1965), Hoel & Levine (1964), Levine (1966), Elfving
(1958) and Murty (1971) are all very specific to their problems and do not extend to the
general linear regression models.

3. E-optimal Designs

Ehrenfeld (1955) was an early advocate of using an E-optimal design. He called a
design $\xi^*$ E-optimal if

$$
\xi^* = \arg \min_{\xi \in \Xi} \max_{c \in \Theta} c^T M^{-1}(\xi)c,
$$

where $\Theta = \{c \in R^k | c^T c = 1\}$. An important property possessed by such a design is that it maximizes the minimum power of a certain $F$-test for testing a subset of $\beta$'s equal to zero. It appears that neither large-sample nor small-sample $E$-optimal regression designs have been intensively studied, perhaps primarily because of the mathematical difficulty involved. For example, there does not seem to exist a practical way to verify if a large-sample design is $E$-optimal, let alone an algorithm for generating $E$-optimal designs in a general regression setup. In this section, we show that new insight into $E$-optimal designs can be gained from using (2.4).

Let $\gamma_{\max}(M^{-1}(\xi))$ denote the maximum eigenvalue of the matrix $M^{-1}(\xi)$ for a design $\xi$ and let $M^{-1}(\xi)$ have corresponding normalized row eigenvectors $Q(\xi)$. Kiefer (1974) showed that, for homoscedastic models, a design $\xi^*$ is $E$-optimal if and only if

$$
\max_{\xi \in \Xi} \gamma_{\max}(Q(\xi^*)M^{-1}(\xi^*)M^{-1}(\xi^*)Q(\xi^*)^T) = \gamma_{\max}(M^{-1}(\xi^*)).
$$

(3.1)

The above equivalence condition is interesting from a theoretical standpoint but it is not clear how it can be used to construct an $E$-optimal design. If we set $\Psi(c, M(\xi)) = \text{tr}\{M^{-1}(\xi)cc^T\}$ and $Z = \Theta$ in (2.4), it follows that

$$
A(\xi) = \{a \in \Theta | a^T M^{-1}(\xi)a = \max_{y \in \Theta} y^T M^{-1}(\xi)y\}
$$

is the set of normalized eigenvectors corresponding to $\gamma_{\max}(M^{-1}(\xi))$. It is straightforward to check that the hypotheses for (2.4) are satisfied and so a design $\xi^*$ is $E$-optimal if and only if there exists a probability measure $\mu^*$ on $A(\xi^*)$ such that, for all $x \in \Omega$,

$$
\int_{A(\xi^*)} u^T M^{-1}(\xi^*)f(x)f^T(x)M^{-1}(\xi^*)u\mu^*(du) - a^T M^{-1}(\xi^*)a \leq 0.
$$

(3.2)

We prefer the equivalence statement in (3.2) to (3.1) because it is easier to interpret. In (3.1), it is usually not possible to check the validity of the equation since we have to test for each design $\xi$ in $\Xi$. On the other hand, for a given $\xi$, (3.2) can be used to verify its optimality by integrating the integrand on the left-hand side of (3.2) with respect to any $A$-measure for $\xi$ and observing if (3.2) holds. A further advantage of (3.2) is that it offers a hint for searching for an $E$-optimal design when the current design is not optimal; see § 5.

Notice that in (3.2) the support points of $\mu^*$ cannot exceed the geometric multiplicity of $\gamma_{\max}(M^{-1}(\xi^*))$. Thus, if we assume the arithmetic multiplicity of $\gamma_{\max}(M^{-1}(\xi^*))$ is one, that is $\gamma_{\max}(M^{-1}(\xi^*))$ is simple, then the geometric multiplicity of $\gamma_{\max}(M^{-1}(\xi^*))$ must equal one as well. This follows from the fact that the geometric multiplicity of any eigenvalue is less than or equal to its arithmetic multiplicity. Together this implies that $\mu^*$ is degenerate and the equivalence condition (3.2) simplifies to

$$
\{a^T M^{-1}(\xi^*)f(x)\}^2 - \gamma_{\max}(M^{-1}(\xi^*)) \leq 0.
$$

(3.3)
This is precisely the same condition imposed by Kiefer (1974, Theorem 6) in order to arrive at the simpler condition (3.3). The differentiability argument used by Kiefer is more involved than the argument used here. We give an example now showing that the \( \mu^* \) in Theorem 2.1 need not be unique.

**Example 3.1.** Consider finding an \( E \)-optimal design for (i) \( p = 2 \) and (ii) \( p = 3 \). For (i), a natural guess is the design \( \xi^* \), which places equal mass at \( \pm 1 \). Clearly,

\[
A(\xi^*) = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}
\]

and a short calculation shows \( \mu^* \) can be taken to be any probability measure supported on \( A(\xi^*) \). For (ii), the optimality of the design \( \xi^* = 0.2\delta_{-1} + 0.6\delta_0 + 0.2\delta_1 \) is verified by noting that the largest eigenvalue of \( M^{-1}(\xi^*) \) is 5. The unique normalized eigenvector corresponding to 5 is \( a^T = (1, 0, -2)/5 \). Since \( \mu^* \) is degenerate at the point \( a \), we may apply (3.3) and after reduction obtain, for all \( x \in \Omega \),

\[
(1 - 2x^2)^2 - 1 \leq 0.
\]

This verifies the \( E \)-optimality of \( \xi^* \).

4. **Further Applications**

4.1. **Minimax variance optimal designs**

Let \( Z \) be an arbitrary compact subspace and set \( \Psi(x, M(\xi)) = \text{tr} \{ B(x)M^{-1}(\xi) \} \) with \( B(x) = f(x)f^T(x) \). Consider finding a design \( \xi^* \) so that

\[
\xi^* = \arg\min_{\xi \in \Xi} \max_{x \in Z} d(x, \xi).
\]

The design \( \xi^* \) in (4.1) is the optimal design for minimizing the maximum variance across the compact set \( Z \) and we call such a design a minimax variance optimal design. Again, it can be shown that (2.4) is applicable and, in particular, could be specialized to yield, for example, the Allocation Rule 1 given by Gaylor & Sweeny (1965) and the results of Kiefer & Wolfowitz (1965), Hoel & Levine (1964) and Levine (1966). The latter authors showed that when \( \Omega = [-1, 1] \) the optimal design is \( \xi_{t,p} \) for \( p \geq 2 \) if \( Z = [1, t] \) and \( t > 1 \) or \( Z = [-1, t] \) and \( t \geq 2 \). The calculation in these papers is somewhat complicated; our approach is more direct. The remaining case when \( Z = [-1, t] \) and \(-1 < t \leq 2 \) can be readily solved by appealing to (2.4); the optimal design is

\[
\{(t + 1)/4\} \delta_1 + \{(3 - t)/4\} \delta_{-1}
\]

with \( \mu^* = \{(t - 2)/(t - 3)\} \delta_{-1} + \{1/(3 - t)\} \delta_1 \).

4.2. **Minimax designs with respect to the single parameters**

Here we let \( Z \) be the set containing the standard basis vectors in \( \Omega \). This choice for \( Z \) may be especially useful when the experimenter does not know in advance which parameters in the model are going to be of interest. Elfving (1958, p. 66) proposed this criterion and called it minimax with respect to the single parameters. This criterion assumes all the parameters have been appropriately scaled; otherwise, this criterion is not meaningful. Some sufficient conditions for a design to be minimax with respect to
the single parameters were given by Elfving (1958, p. 66) and Murty (1971) but their assumptions and results are restrictive. The condition (2.4) is more useful in practice. As an illustration, Murty's result that \( \xi^* = \frac{1}{2} \delta_{-1} + \frac{1}{3} \delta_{-0} + \frac{1}{4} \delta_{0} + \frac{1}{5} \delta_{1} \) is minimax with respect to the single parameters for the case \( p = 4 \) is verified by noting that \( A(\xi^*) = \{(0, 0, 0, 1)\} \), and (2.4) yields, after reduction,

\[
x^2(4x^2 - 3)^2 - 1 \leq 0.
\]

This inequality is true for all \( x \in [-1, 1] \) and \( \xi^* \) is minimax with respect to the single parameters.

4.3. \( P \)-optimal designs

For prediction purposes, we may want to seek a design \( \xi^* \) such that

\[
\xi^* = \arg \min_{\xi \in \Xi} \max_{x \in \Omega} \{d(x, \xi) + 1/\lambda(x)\}.
\]  

We shall call such a design a \( P \)-optimal design. Note that the quantity in braces is proportional to the predictive variance at the point \( x \). This criterion is meaningful when our goal is to select a design to predict at some point \( x \) but there is uncertainty as to where \( x \) might be. A \( P \)-optimal design seeks to minimize the maximum predictive variance across the design space, thus giving the experimenter some global protection.

Although (4.2) does not conform to the form (2.3), the mathematical structure for both the design problems is strikingly similar. With the aid of (2.4), it is straightforward to show a design \( \xi^* \) is \( P \)-optimal if and only if there exists a probability measure \( \mu^* \) on \( A(\xi^*) \) such that, for all \( x \) in \( \Omega \),

\[
\int_{A(\xi^*)} \{1/\lambda(u) + \lambda(x)d^2(x, u, \xi^*)\} \mu^*(du) - d(a, \xi^*) - 1/\lambda(a) \leq 0.
\]  

Here,

\[
A(\xi) = \{a \in \Omega | d(a, \xi) + 1/\lambda(a) = \max_{x \in \Omega} \{d(x, \xi) + 1/\lambda(x)\}\}
\]

and \( d(x, u, \xi) \) is as defined near (2.5). The rest follows similarly as in the above cases.

5. Algorithms for minimax optimal designs

In this section we describe algorithms for generating the minimax designs considered in this paper. The ideas behind these algorithms are based on heuristics and are modifications of the iterative scheme of Wong & Cook (1992). For the most part, it is quite obvious how to do this; so we shall be brief and consider only exemplary cases where we are interested in finding \( P \)- and \( E \)-optimal designs.

The close resemblance in the mathematical structure of the problems for finding \( P \)- and \( G \)-optimal designs suggests the following strategy: simply replace \( d(x, \xi) \) by \( d(x, \xi) + 1/\lambda(x) \) everywhere in the description of the algorithm of Wong & Cook (1992) for finding heteroscedastic \( G \)-optimal designs.

Example 5.1. Suppose we want to find the \( P \)-optimal design for \( p = 2 \) and \( \lambda(x) = 1 + x + x^2 \). The modified algorithm was run on a Vax 11/75 computer and, after 500 iterations, we arrived at the design \( \xi = 0.178 \delta_{1} + 0.822 \delta_{-1} \) with \( \mu_{A, \xi} = 0.123 \delta_{1} + 0.877 \delta_{-1} \). After reduction, the left-hand side of (4.3) becomes

\[
P(x) = -0.865 - 0.001x + 0.431x^2 - 0.001x^3 + 0.432x^4,
\]
which is negative for all \( x \in (-1, 1] \). In addition, \( P(-1) = 0 \) and \( P(1) = -0.004 \); thus verifying \( \xi \) is close to the \( P \)-optimal design. The proximity between \( \xi \) and the \( P \)-optimal design can be quantified in terms of the design efficiency as described in my dissertation. For this problem, the \( P \)-optimal design is

\[
\rho = (3 - \sqrt{7})/2 = 0.177, \quad \mu_{A,\xi^*} = 0.122\delta_1 + 0.878\delta_{-1}.
\]

For heteroscedastic models where an \( E \)-optimal design is desired, the modified algorithm is as follows.

**Step 0.** Set \( k = 0 \) and generate a nonsingular starting design \( \xi_k \).

**Step 1.** Determine the set \( A(\xi_k) = \{u_1, u_2, \ldots, u_{n_k}\} \) by finding the normalized eigenvectors corresponding to \( \gamma_{\text{max}}(M^{-1}(\xi_k)) \).

**Step 2.** Find a measure \( \mu_k \) such that

\[
\mu_k = \arg \min_{\mu \in \Xi_{\xi_k}} \max_{x \in \Omega} \lambda(x) \int_{A(\xi_k)} uM^{-1}(\xi_k)f(x)^Tf(x)M^{-1}(\xi_k)u\mu(du),
\]

where \( \Xi_{\xi_k} \) is the set of all probability measures with support at \( A(\xi_k) \).

**Step 3.** Check if

\[
\max_{x \in \Omega} \lambda(x) \int_{A(\xi_k)} u^TM^{-1}(\xi_k)f(x)^Tf(x)M^{-1}(\xi_k)u\mu_k(du) - \varepsilon \leq 0.
\]

where \( \varepsilon \) is a pre-assigned small positive number of order, say, 0.001. If yes, stop and conclude that \( \xi_k \) is \( \varepsilon \) \( E \)-optimal for the problem. Otherwise, proceed to Step 4.

**Step 4.** A new point, \( x_k \), is added to the current design \( \xi_k \). Following standard optimal design strategies, \( x_k \) is chosen so that

\[
\int_{A(\xi_k)} \lambda(x_k)u^TM^{-1}(\xi_k)f(x_k)f^T(x_k)M^{-1}(\xi_k)u\mu_k(du) = \arg \max_{x \in \Omega} \int_{A(\xi_k)} \lambda(x)u^TM^{-1}(\xi_k)f(x)f^T(x)M^{-1}(\xi_k)u\mu_k(du).
\]

This gives a new design \( \xi_{k+1} = (1 - \alpha_k)\xi_k + \alpha_k\delta_{x_k} \), where \( \alpha_k \) is some suitable weighting sequence satisfying \( \sum \alpha_k < \infty \) and \( \alpha_k \to 0 \).

**Step 5.** Set \( k \) to \( k+1 \) and go to Step 1 with the new design \( \xi_{k+1} \).

This shows how to modify an existing design to get closer to the optimal design in general. It is known that the algorithm does not have the monotonicity property; it does not guarantee that the augmented design will have a smaller criterion value than the existing one. Our experience is that the algorithm always generates designs with sufficiently high efficiencies for all the problems we have looked at. This algorithm appears promising and is currently under further refinement to improve its efficiency.

**Acknowledgements**

Much of this paper is taken from the author's Ph.D. thesis, written under the supervision of R. Dennis Cook at the University of Minnesota. I am also thankful to A. A. Afifi for reading a previous version of this paper.
REFERENCES


[Received March 1991. Revised December 1991]