The Annals of Statistics
1976, Vol. 4, No. 6, 1271–1279

REMEZ'S PROCEDURE FOR FINDING OPTIMAL DESIGNS

BY WILLIAM J. STUDDEN AND JIA-YEONG TSAY

Purdue University and University of Cincinnati

The Remez exchange procedures of approximation theory are used to find the optimal design for the problem of estimating $c'\theta$ in the regression model $EY(x) = \theta_1 f_1(x) + \theta_2 f_2(x) + \cdots + \theta_k f_k(x)$, when c is not a linear combination of less than k vectors of the form $f(x)$. A geometric approach is given first with a proof of convergence. When the design space is a closed interval, the Remez exchange procedure is illustrated by two examples. This type procedure can be used to find the optimal design very efficiently, if there exists an optimal design with k support points.

1. Introduction. Let $f' = (f_1, f_2, \cdots, f_k)$ be a vector of linearly independent continuous functions on a compact set $\mathscr{X}$. For each x or "level" in $\mathscr{X}$ an experiment can be performed whose outcome is a random variable $Y(x)$ with mean value $\theta' f(x) = \sum \theta_i f_i(x)$ and variance σ^2, independent of x. The functions f_i, $i = 1, 2, \cdots, k$ are called the regression functions and assumed known to the experimenter while the vector of parameters $\theta' = (\theta_1, \theta_2, \cdots, \theta_k)$ and σ^2 are unknown.

An experimental design is a probability measure ξ on $\mathscr{X}$. If the unknown parameter vector θ is estimated by the method of least squares with N uncorrelated observations $\{Y(x_i)\}_{i=1}^N$, then the covariance matrix of the best linear unbiased estimator $\hat{\theta}$ is given by $\sigma^2/N \cdot M^{-1}(\xi)$, where $M(\xi) = (m_{ij}(\xi))$, with $m_{ij}(\xi) = \int f_i(x) f_j(x)\, d\xi(x)$, is the information matrix of the design ξ.

A fairly general problem in design theory is to minimize a convex function $\Psi(M(\xi))$ of the information matrix $M(\xi)$. For example, $\Psi(M(\xi)) = \operatorname{tr} BM^{-1}(\xi)$ for B positive semidefinite (L-optimality) or $\Psi(M(\xi)) = -\log |M(\xi)|$ where $|M(\xi)|$ denotes the determinant of $M(\xi)$ (D-optimality). Recently a number of equivalence theorems and closely related iterative procedures have appeared for minimizing $\Psi(M(\xi))$, see Kiefer (1974) for references. The purpose of this paper is to describe and study some special iterative procedures which can be used to find optimal designs very efficiently, if there exists an optimal design with k support points. In Section 2 we provide and discuss the procedures. Section 3 contains a special case which in approximation theory is called the Remez exchange procedure. Examples are given in Section 4.

2. Procedures. One of the general iterative procedures for minimizing $\Psi(M)$

Received August 1974; revised December 1975.

[1] This research was supported by the National Science Foundation Grant 3352x2 and the David Ross Grant 6633. Reproduction is permitted in whole or in part for any purposes of the United States Government.

AMS 1970 *subject classification.* Primary 62K05.

Key words and phrases. Optimal design, information matrix, Remez exchange procedures, regression functions.

is the following: if at the nth step we are at $M(\xi_n) = M_n$ we then move locally in a direction with "steepest descent." This is, we choose an information matrix $\bar{M}_n$ so that

$$g(\alpha) = \Psi((1-\alpha)M_n + \alpha\bar{M}_n)$$

has a minimum derivative at $\alpha = 0$. Then let $M_{n+1} = (1-\alpha)M_n + \alpha\bar{M}_n$ and α be suitably chosen to give a decrease in Ψ. Since the set of all information matrices $M(\xi)$ is "spanned" by the set of matrices $M(\xi_x) = f(x)f'(x)$, $x \in \mathscr{X}$ (ξ_x is a design with mass one at the point x) we restrict the matrices $\bar{M}_n$ to be of the form $\bar{M} = f(x)f'(x)$ and choose the point x minimizing $g'(0)$ to have $M_{n+1} = (1-\alpha)M_n + \alpha f(x)f'(x)$. It is known that $g'(0) = \operatorname{tr} \nabla\Psi(M_n)\cdot(f(x)f'(x) - M_n)$ where $\nabla\Psi(M)$ is the $k \times k$ matrix with entries $(\nabla\Psi(M))_{ij} = (\partial/\partial m_{ij})\Psi(M)$, and Ψ is convex, thus for any design ξ_n,

$$\text{(2.1)} \qquad \inf_x \operatorname{tr} \nabla\Psi(M_n)\cdot(f(x)f'(x) - M_n) \leq 0\,, \qquad \text{or equivalently}$$
$$\inf_x f'(x)\nabla\Psi(M_n)f(x) \leq \operatorname{tr} \nabla\Psi(M_n)M_n\,.$$

In certain special cases of Ψ, the $\alpha = \alpha_n$ at the nth step may be explicitly chosen in an optimal manner. The Fedorov procedure belongs to this type. The simplest method to choose α_n is just to use any $\alpha_n \to 0$ to obtain some sort of convergence, and $\sum \alpha_n = \infty$ to prevent convergence before reaching a minimum. Wynn's procedure (1970, 1972) is an example.

If there exists an optimal design with k support points, a special iterative procedure can be used, which is described as follows. Given a set S_0 of k points $x_{01}, x_{02}, \cdots, x_{0k}$ such that the k vectors $f(x_{01}), f(x_{02}), \cdots, f(x_{0k})$ are linearly independent we find the optimal weights p_{0i} on x_{0i}, $i = 1, 2, \cdots, k$, that is,

$$\xi_0 = \begin{Bmatrix} x_{01}, x_{02}, \cdots, x_{0k} \\ p_{01}, p_{02}, \cdots, p_{0k} \end{Bmatrix}$$

is an optimal design on S_0. Then we find a point x_0 minimizing $f'(x)\nabla\Psi(M(\xi_0))f(x)$ and exchange x_0 with a point x_{0i} in S_0 according to some exchange rules to form a new set S_1 of k points $x_{11}, x_{12}, \cdots, x_{1k}$. Repeating the same procedure on S_1 as S_0, we get an optimal design ξ_1 on S_1 and point x_1 minimizing $f'(x)\nabla\Psi(M(\xi_1))f(x)$ which is used to exchange with some point x_{1i} in S_1 to form S_2. Continuing this process, we obtain a sequence of designs $\{\xi_n\}$ which converges very quickly to the optimal design, sometimes even in a few steps. If the set S_0 happens to be the support of the optimal design, then we obtain the optimal design in one step only.

Now the technical problems remaining are how to find the optimal weights and how to exchange the points at each step. These problems depend on the optimality criteria. For example, in the case of D-optimality we may take $1/k$ for all points in S_n and exchange x_n with the point closest to it in S_n or by some other means. The explicit forms for finding optimal weights and exchanging points will be given with a proof of convergence for the following special L-optimality.

Let c be a vector which cannot be written as a linear combination of less than k vectors in the set $\{f(x) \mid x \in \mathscr{X}\}$, and let $\Psi(M(\xi)) = c'M^{-1}(\xi)c$ (c-optimality). Then $\nabla\Psi(M(\xi)) = -M^{-1}(\xi)cc'M^{-1}(\xi)$, hence from (2.1) we have

$$\sup_x (c'M^{-1}(\xi)f(x))^2 \geqq c'M^{-1}(\xi)c \tag{2.2}$$

where the equality holds if and only if ξ is c-optimal by Theorem 2.9.2 of Fedorov (1972). The existence of a c-optimal design with k support points is assured by a theorem of Elfving (see Karlin and Studden 1966)) which states that ξ^* is c-optimal if and only if there exists a function ϕ with $|\phi(x)| = 1$ such that $\int \phi(x)f(x)\,d\xi^*(x) = \beta_* c$ for $\beta_*^{-2} = \inf_\xi c'M^{-1}(\xi)c$ and $\beta_* c$ is a boundary point of a set R which is the convex hull of the set $\{\pm f(x) \mid x \in \mathscr{X}\}$.

From the relation $\beta_*^{-2} = \inf_\xi c'M^{-}(\xi)c$, it follows that over the set of vectors d with $d'M(\xi)d \neq 0$

$$\begin{aligned} \beta_*^{-2} &= \inf_\xi \sup_d \frac{(c'd)^2}{d'M(\xi)d} \\ &\geqq \inf_\xi \frac{(c'd)^2}{\int (d'f(x))^2\,d\xi(x)} \\ &= \frac{(c'd)^2}{\sup_x (d'f(x))^2}\,. \end{aligned} \tag{2.3}$$

Thus for any vector d such that $c'd = 1$, we have

$$\sup_x |d'f(x)| \geqq \beta_*\,. \tag{2.4}$$

It is interesting to note that from (2.3) and (2.4) the design problem (minimizing $c'M^{-1}(\xi)c$ over the set of designs ξ) becomes an approximation problem (minimizing $\sup_x |d'f(x)|$ over the set of vectors d with $c'd = 1$).

Let $S_0 = \{x_{01}, x_{02}, \cdots, x_{0k}\}$ such that the vectors $f(x_{0i})$, $i = 1, 2, \cdots, k$ are linearly independent. By the Elfving theorem mentioned above applied to the set S_0, the optimal weights p_{0i} on x_{0i} are a solution of the equation

$$\textstyle\sum_{i=1}^{k} \phi_{0i} p_{0i} f(x_{0i}) = \beta_0 c \tag{2.5}$$

where $\phi_{0i} = \pm 1$, $p_{0i} \geqq 0$ and $\sum p_{0i} = 1$, $\beta_0 > 0$ and β_0^{-2} is the minimum value of $c'M^{-1}(\xi)c$ for ξ with support on S_0. Hence

$$\xi_0 = \begin{Bmatrix} x_{01}, x_{02}, \cdots, x_{0k} \\ p_{01}, p_{02}, \cdots, p_{0k} \end{Bmatrix}$$

is a c-optimal design on S_0 and the vector $d_0' = c'M^{-1}(\xi_0)/c'M^{-1}(\xi_0)c$ minimizes the value $\sup_{x\in S_0} |d'f(x)|$. Now the vector d_0 gives a hyperplane $d_0'z = \beta_0$ ($z' = (z_1, z_2, \cdots, z_k)$) at $\beta_0 c$ to the set R_0 defined as the convex hull of the set $\{\pm f(x) \mid x \in S_0\}$. Let $\varphi_0(x) = d_0'f(x)$. From (2.5) $\beta_0 c$ is a convex combination of $\phi_{0i} f(x_{0i})$ and each $\phi_{0i} f(x_{0i})$ lies on the hyperplane $d_0'z = \beta_0$, that is, $\phi_{0i}\varphi_0(x_{0i}) = \beta_0$ for $i = 1, 2, \cdots, k$. If ξ_0 is c-optimal then we stop, in this case $\sup_x |\varphi_0(x)| = \beta_0$. If φ_0 is not, we then choose the point x_0 maximizing $|\varphi_0(x)|$ so that $\phi_0\varphi_0(x_0) > \beta_0$, or equivalently we find a vector $\phi_0 f(x_0)$ which lies on the side of the hyperplane $d_0'z = \beta_0$ opposite the origin and farthest from the hyperplane. If we can exchange $\phi_0 f(x_0)$ with one of the vectors $\phi_{0i} f(x_{0i})$, say $\phi_{0j} f(x_{0j})$, so that $\beta_1 c$ ($\beta_1 > 0$) is a convex combination of the new set of vectors, then $\beta_1 \geqq \beta_0$. This is true, since

$$
\begin{aligned}
\beta_1 &= \beta_1 d_0'c = d_0'\beta_1 c \\
&= d_0'[\textstyle\sum_{i \neq j} p_{1i}\phi_{0i} f(x_{0i}) + p_{1j}\phi_0 f(x_0)] \\
&= d_0'[\textstyle\sum_i p_{1i}\phi_{0i} f(x_{0i}) + p_{1j}(\phi_0 f(x_0) - \phi_{0j} f(x_{0j}))] \\
&= \textstyle\sum_i p_{1i}\phi_{0i}\varphi_0(x_{0i}) + p_{1j}(\phi_0\varphi_0(x_0) - \phi_{0j}\varphi_0(x_{0j})) \\
&= \beta_0 + p_{1j}(|\varphi_0(x_0)| - \beta_0) \,.
\end{aligned} \tag{2.6}
$$

In order to determine how the exchange should be made we let, for the convenience of notations, $a = \phi_0 f(x_0)$, $a_i = \phi_{0i} f(x_{0i})$, $p_i = p_{0i}$ for $i = 1, 2, \cdots, k$. Then (2.5) can be written as

$$\beta_0 c = \textstyle\sum_{i=1}^k p_i a_i \tag{2.7}$$

and we wish an exchange so that a similar equation holds. We simply take a representation

$$a = \textstyle\sum q_i a_i \tag{2.8}$$

and consider an exchange using a_j with $q_j \neq 0$. Solving (2.8) for a_j and substituting in (2.7) we get

$$
\begin{aligned}
\beta_0 c &= \textstyle\sum_{i \neq j} p_i a_i + p_j(a - \sum_{i \neq j} q_i a_i)/q_j \\
&= \textstyle\sum_{i \neq j} q_i \left(\frac{p_i}{q_i} - \frac{p_j}{q_j}\right) a_i + \frac{p_j}{q_j}\, a \,.
\end{aligned} \tag{2.9}
$$

In order to have all coefficients positive we choose $j = j_0$ to give minimum value for p_j/q_j in those $q_j > 0$. Let $s = \sum_{i \neq j_0} q_i(p_i/q_i - p_j/q_j) + p_{j_0}/q_{j_0}$, then a renormalization of (2.9) gives

$$\beta_1 c = \textstyle\sum p_i' a_i' \tag{2.10}$$

where $p_i' = q_i(p_i/q_i - p_{j_0}/q_{j_0})/s$, $a_i' = a_i$ for $i \neq j_0$ and $p_{j_0}' = p_{j_0}/sq_{j_0}$, $a_{j_0}' = a$, and $\beta_1 = \beta_0/s$.

The procedure above holds at each step, hence in (2.5) and (2.6) replacing 0 and 1 with n and $n + 1$ we have

$$\beta_n c = \textstyle\sum_{i=1}^k p_{ni}\phi_{ni} f(x_{ni}) \,, \qquad \text{and} \tag{2.11}$$

$$\beta_{n+1} = \beta_n + p_{(n+1)j}(|\varphi_n(x_n)| - \beta_n) \,, \tag{2.12}$$

from which we obtain a sequence of designs $\{\xi_n\}$ with

$$\xi_n = \begin{Bmatrix} x_{n1}, x_{n2}, \cdots, x_{nk} \\ p_{n1}, p_{n2}, \cdots, p_{nk} \end{Bmatrix}$$

to be c-optimal on $S_n = \{x_{n1}, x_{n2}, \cdots, x_{nk}\}$ and

$$|\varphi_n(x_n)| = \sup_{x \in S_n} |c'M^{-1}(\xi_n) f(x)/c'M^{-1}(\xi_n)c| \,.$$

THEOREM. *For the sequence of designs* $\{\xi_n\}$ *generated above,* $c'M^{-1}(\xi_n)c$ *converges monotonically to* $c'M^{-1}(\xi^*)c$ *where* ξ^* *is c-optimal.*

PROOF. Since $c'M^{-1}(\xi_n)c = \beta_n^{-2}$ and $c'M^{-1}(\xi^*)c = \beta_*^{-2}$, we will show that β_n converges monotonically to β_*.

In (2.12), $|\varphi_n(x_n)| - \beta_n \geqq 0$ for every n, it follows that

$$\beta_0 \leqq \beta_1 \leqq \beta_2 \leqq \cdots \leqq \beta_* .$$

Hence β_n converges monotonically, say, to $\bar{\beta}$. To prove $\bar{\beta} = \beta_*$ it suffices to show that the limit inf

$$\liminf_n p_{ni} > 0 \qquad \text{for all} \quad i = 1, 2, \cdots, k , \tag{2.13}$$

since the convergence of β_n and (2.12) imply $|\varphi_n(x_n)| - \beta_n \to 0$, but $\beta_n \leqq \bar{\beta} \leqq \beta_* \leqq |\varphi_n(x_n)|$, thus $|\varphi_n(x_n)|$ and β_n converges to the same limit $\bar{\beta} = \beta_*$.

Suppose (2.13) is not true, then there exists an i_0 and a subsequence $\{p_{n_j i_0}\}$ such that $p_{n_j i_0} \to 0$. We refine the subsequence so that all $p_{n_j i}$, $x_{n_j i}$, and $\phi_{n_j i}$ converge, say, to $\bar{p}_i$, $\bar{x}_i$, $\bar{\phi}_i$ for $i = 1, 2, \cdots, k$ where $\bar{p}_{i_0} = 0$. From (2.11) we have

$$\beta_{n_j} c = \sum_{i=1}^{k} p_{n_j i} \phi_{n_j i} f(x_{n_j i}) .$$

Let $n_j \to \infty$, it follows by continuity that

$$\begin{aligned} \bar{\beta} c &= \sum_{i=1}^{k} \bar{p}_i \bar{\phi}_i f(\bar{x}_i) \\ &= \sum_{i \neq i_0}^{k} \bar{p}_i \bar{\phi}_i f(\bar{x}_i) . \end{aligned}$$

But $\bar{\beta} > 0$, hence c is a linear combination of $k - 1$ vectors in $\{f(x) \mid x \in \mathscr{X}\}$ which contradicts the assumption of c. This proves the theorem.

REMARKS. The procedure described above is closely related to the Silvey and Titterington procedure (1973). The latter at each step has the best k-point subset with optimal weights on it. The former, however, does not generally have the best k-point subset, although the weights are also optimal on the set. Each has its own advantage. For instance, at the nth step ξ_n is an optimal design on the set $S_n = \{x_{n1}, x_{n2}, \cdots, x_{nk}\}$ and x_n is the point maximizing $|\varphi_n(x)|$. In the Silvey and Titterington procedure ξ_{n+1} is chosen to be the best k-point design (in terms of β_{n+1}) from the set $S_n^* = \{x_n, x_{n1}, x_{n2}, \cdots, x_{nk}\}$. Therefore, in order to determine the best β_{n+1} among all k-point subsets of S_n^* we have to solve the system of equations (2.11) k times. But in the procedure given above, x_n is exchanged with some point in S_n according to certain rules. Thus S_{n+1} is not generally the best k-point design from S_n^*, but we only have to solve (2.11) once instead of k times. Of course, the best advantage of the Silvey and Titterington procedure is no restriction on the number of support points for the optimal design, hence its applicability is wider. But the complicated calculation is the price to be paid.

3. A special case. For the procedure given in Section 2 if $\mathscr{X} = [a, b]$, there is a special exchange method which in approximation theory is called the Remez type procedure or the Remez exchange procedure. This procedure is taken from Meinardus (1967), but in a design theory context.

As before, c is assumed to be a vector which can not be written as a linear combination of less than k vectors in $\{f(x) \mid x \in [a, b]\}$. Therefore, the determinants

$$D_i(c) = D_i(c; x_1, x_2, \cdots, x_k) = |f(x_1), \cdots, f(x_{i-1}), c, f(x_{i+1}), \cdots, f(x_k)|$$

are never zero and they alternate in sign for any set of k points $a \leqq x_1 < x_2 < \cdots < x_k \leqq b$. Now we start with a set $S_0 = \{x_{01}, x_{02}, \cdots x_{0k}\}$ such that $a \leqq x_{01} < x_{02} < \cdots < x_{0k} \leqq b$. Solving (2.5) we have

$$p_{0i} = |D_i(c)|/\sum_{j=1}^{k} |D_j(c)|\,, \qquad \phi_{0i} = \operatorname{sgn} D_i(c)\,, \quad \text{and} \qquad \beta_0^{-2}$$

is the minimum value of $c'M^{-1}(\xi)c$ for ξ with support on S_0. Let ξ_0 denote the above design and let the function $\varphi_0(x) = c'M^{-1}(\xi_0)f(x)/c'M^{-1}(\xi_0)c$. Then we choose a new set S_1 of k points $a \leqq x_{11} < x_{12} < \cdots < x_{1k} \leqq b$ such that

(i) $|\varphi_0(x_{1i})| \geqq \beta_0$, $i = 1, 2, \cdots, k$
(ii) $|\varphi_0(x_{1i_0})| > \beta_0$ for some i_0
(iii) $\operatorname{sgn} \varphi_0(x_{1i}) = \alpha \operatorname{sgn} \varphi_0(x_{0i})$ where the constant α is either 1 or -1.

Continuing the same procedure on S_1, then S_2 and so on, we obtain a sequence of the designs $\{\xi_n\}$ with $\beta_n^{-2} = c'M^{-1}(\xi_n)c$ which can be shown converging monotonically to $\inf_\xi c'M^{-1}(\xi)c$ by the same proof of the theorem in Section 2.

With regard to the conditions (i), (ii) and (iii) for the new set of points there are two usual methods of proceeding. Typically the function $\varphi_0(x)$ will have k-2 local extrema x_{1i}, $i = 2, 3, \cdots, k-1$ and we use these together with $x_{11} = a$ and $x_{1k} = b$. The other method is to just choose ω to give $|\varphi_0(\omega)| = \max_x |\varphi_0(x)|$ and then exchange ω with one of the x_{0i} values to satisfy (iii). Roughly speaking, this entails exchanging ω with an adjacent x_{0i} value (if we connect two ends a and b) for which φ_0 has the same sign. Note that the values $\varphi_0(x_{0i})$, $i = 1, 2, \cdots, k$ also alternate in sign. In general we use the following rule:

ω value	$\operatorname{sgn} \varphi_0(\omega) =$	ω replaces
$a \leqq \omega < x_{01}$	$\operatorname{sgn} \varphi_0(x_{01})$	x_{01}
$a \leqq \omega < x_{01}$	$-\operatorname{sgn} \varphi_0(x_{01})$	x_{0k}
$1 \leqq i \leqq k-1$		
$x_{0i} < \omega < x_{0i+1}$	$\operatorname{sgn} \varphi_0(x_{0i})$	x_{0i}
$x_{0i} < \omega < x_{0i+1}$	$-\operatorname{sgn} \varphi_0(x_{0i})$	x_{0i+1}
$x_{0k} < \omega \leqq b$	$\operatorname{sgn} \varphi_0(x_{0k})$	x_{0k}
$x_{0k} < \omega \leqq b$	$-\operatorname{sgn} \varphi_0(x_{0k})$	x_{01}

In the next section, two examples will be given to illustrate this exchange procedure.

4. Examples.

EXAMPLE 1. Let $\mathscr{X} = [-1, 1]$, $f'(x) = (1, x)$ and $c' = (0, 1)$. For the initial set $S_0 = \{x_{01}, x_{02}\}$ we use $x_{01} = -\frac{1}{2}$ and $x_{02} = \frac{3}{4}$. Then from (2.5),

$$\xi_0 = \begin{Bmatrix} -\frac{1}{2}, & \frac{3}{4} \\ \frac{1}{2}, & \frac{1}{2} \end{Bmatrix}, \qquad \varphi_0(x) = x - \tfrac{1}{8} \qquad \text{and} \qquad \omega = -1$$

giving $|\varphi_0(\omega)| = \max_x |\varphi_0(x)|$. Moreover $\phi_{01} = -1 = \operatorname{sgn} \varphi_0(x_{01})$, $\phi_{02} = 1 = \operatorname{sgn} \varphi_0(x_{02})$ and $\beta_0 = \frac{5}{8}$. One can easily show that $\omega = -1$ must be exchanged with $x_{01} = -\frac{1}{2}$ giving $\beta_1 = \frac{7}{8}$. The exchange with x_{02} will give a decrease to $\beta_1 = \frac{1}{4}$ which is not what we want. The next step will produce $\omega = 1$ which will exchanged with $\frac{3}{4}$. Therefore the c-optimal design

$$\xi^* = \begin{Bmatrix} -1, & 1 \\ \frac{1}{2}, & \frac{1}{2} \end{Bmatrix}$$

is obtained in two steps.

EXAMPLE 2. Let $f'(x) = (1, x, x^2, (x - \eta)_+{}^2)$ on $\mathscr{X} = [-1, 1]$, where $(x - \eta)_+{}^2 = (x - \eta)^2$ if $x \geqq \eta$ and equals zero for $x < \eta$. We consider the cases $c' = (0, 0, 0, 1)$ and $\eta = 0, .4, .8$. Four equally spaced points on $[-1, 1]$ are used for the initial set of points x_{0i}, $i = 1, 2, 3, 4$. The procedure is terminated if the critical value

$$\lambda_n = \frac{||\varphi_n|| - \beta_n}{\beta_n} < 10^{-5}.$$

where $||\varphi_n|| = \max_x |\varphi_n(x)|$. The results can be found in Table 1.

The design ξ_3 for each case is then

Case (i). $\eta = 0$

$$\xi_3 = \begin{Bmatrix} -1 & -.4142 & .4137 & 1 \\ .1465 & .3537 & .3535 & .1463 \end{Bmatrix}$$

and

$$\beta_3^{-2} = 135.8824.$$

TABLE 1

	n	x_{n1}	x_{n2}	x_{n3}	x_{n4}	β_n	λ_n
$\eta = 0$	0	-1	$-.3333$	.3333	1	8.3333×10^{-2}	4.1667×10^{-2}
	1	-1	$-.4166$	.3333	1	8.4641×10^{-2}	3.8339×10^{-2}
	2	-1	$-.4166$	.4137	1	8.5785×10^{-2}	3.5083×10^{-5}
	3	-1	$-.4142$	.4137	1	8.5786×10^{-2}	1.3237×10^{-6}
$\eta = .4$	0	-1	$-.3333$	.3333	1	4.5000×10^{-2}	1.0345×10^{0}
	1	-1	$-.3333$	.5862	1	6.3108×10^{-2}	2.2624×10^{-2}
	2	-1	$-.2545$	.5862	1	6.3514×10^{-2}	7.5706×10^{-4}
	3	-1	$-.2545$	.5941	1	6.3534×10^{-2}	6.3136×10^{-8}
$\eta = .8$	0	-1	$-.3333$	.3333	1	5.0000×10^{-3}	3.9130×10^{0}
	1	-1	$-.3333$	.8261	1	1.3498×10^{-2}	1.5178×10^{-1}
	2	-1	$-.0922$	.8261	1	1.3799×10^{-2}	1.6546×10^{-3}
	3	-1	$-.0922$	.8309	1	1.3810×10^{-2}	1.7458×10^{-8}

Case (ii). $\eta = .4$

$$\xi_3 = \begin{Bmatrix} -1 & -.2545 & .5941 & 1 \\ .0938 & .2810 & .4062 & .2190 \end{Bmatrix}$$

and

$$\beta_3^{-2} = 247.7351\,.$$

Case (iii). $\eta = .8$

$$\xi_3 = \begin{Bmatrix} -1 & -.0922 & .8309 & 1 \\ .0396 & .1437 & .4604 & .3563 \end{Bmatrix}$$

and

$$\beta_3^{-2} = 5243.6836\,.$$

The Fedorov procedure for this example was run for 30 iterations in each case, the rounded-off design (as described in Fedorov (1972), page 109) for each case is

Case (i). $\eta = 0$

$$\tilde{\xi}_{30} = \begin{Bmatrix} -1 & -.3986 & .3958 & 1 \\ .1424 & .3440 & .3712 & .1424 \end{Bmatrix}$$

and

$$c'M^{-1}(\tilde{\xi}_{30})c = 136.3643\,.$$

Case (ii). $\eta = .4$

$$\tilde{\xi}_{30} = \begin{Bmatrix} -1 & -.3166 & .5305 & 1 \\ .1144 & .2427 & .4633 & .1796 \end{Bmatrix}$$

and

$$c'M^{-1}(\tilde{\xi}_{30})c = 267.8787\,.$$

Case (iii). $\eta = .8$

$$\tilde{\xi}_{30} = \begin{Bmatrix} -1 & -.2657 & .7589 & 1 \\ .0677 & .1088 & .4955 & .3280 \end{Bmatrix}$$

and

$$c'M^{-1}(\tilde{\xi}_{30})c = 7267.5000\,.$$

From the comparison above it is known that the Fedorov procedure converges much more slowly than the Remez procedure. The main drawback of the Fedorov procedure is that it cannot throw out bad points during the iterative process. This becomes more serious when the initial design contains some bad points with heavy weights as Case (iii) in this example. Although the Atwood procedure (1973) can overcome this problem, it still cannot give the right point and throw out the bad point simultaneously. Further, it cannot give the optimal weights on the support points at each step. Thus it needs many steps to get the optimal design, even if we use the support points of the optimal design with inadequate weights for the initial design.

Acknowledgments. The authors are grateful to the referees and the Associate Editor for their useful suggestions on the revision of this paper.

REFERENCES

[1] Atwood, C. L. (1973). Sequence converging to D-optimal designs of experiment. *Ann. Statist.* **1** 342–352.

[2] Fedorov, V. V. (1972). *Theory of Optimal Experiments.* Academic Press, New York.

[3] Karlin, S. and Studden, W. J. (1966). Optimal experimental designs. *Ann. Math. Statist.* **37** 783–815.

[4] Kiefer, J. (1974). General equivalence theory for optimal designs (approximation theory). *Ann. Math. Statist.* **2** 849–879.

[5] Meinardus, G. (1967). *Approximation of Functions.* Springer-Verlag, New York.

[6] Silvey, J. D. and Titterington, D. M. (1973). A geometric approach to optimal design theory. *Biometrika* **60** 21–32.

[7] Wynn, H. P. (1970). The sequential generation of D-optimal experimental designs. *Ann. Math. Statist.* **41** 1655–1664.

[8] Wynn, H. P. (1972). Results in the theory and construction of D-optimum experimental designs. *J. Roy. Statist. Soc. Ser. B* **34** 133–147.

Department of Statistics
Purdue University
West Lafayette, Indiana 47907

Department of environmental Health
Division of Biostatistics and Epidemiology
University of Cincinnati Medical College
Cincinnati, Ohio 45267