On Remez Type Procedures for Calculating Optimal Designs*

W. J. Studden

and

Jia-Yeong Tsay

Purdue University

Department of Statistics

Mimeograph Series #326

July 1973

---

*This research was supported by the National Science Foundation Grant 33552x2 and the Davis Ross Grant 6633.

On Remez Type Procedures for

Calculating Optimal Designs*

W. J. Studden

and

Jia-Young Tsay
Purdue University

1. Introduction

Let $f' = (f_0, f_1,\ldots,f_k)$ be a vector of linearly independent continuous functions on a compact set $\mathcal{X}$. For each x or "level" in $\mathcal{X}$ an experiment can be performed whose outcome is a random variable $Y(x)$ with mean value $\theta' f(x) = \Sigma\, \theta_i\, f_i(x)$ and variance σ^2, independent of x. The functions f_i, $i=0,1,\ldots,k$ are called the regression functions and assumed known to the experimenter while the vector of parameters $\theta' = (\theta_0,\theta_1,\ldots,\theta_k)$ and σ^2 are unknown. An experimental design is a probability measure ξ on $\mathcal{X}$. If ξ concentrates mass ξ_i at the points x_i, $i = 1, 2,\ldots,r$ and $\xi_i N = n_i$ are integers, the experimenter takes N uncorrelated observations, n_i at each x_i, $i = 1, 2,\ldots,r$. The covariance matrix of the least squares estimates of the parameters θ_i is then given by $\frac{\sigma^2}{N} M^{-1}(\xi)$ where $M(\xi) = (m_{ij}(\xi))$, $m_{ij}(\xi) = \int f_i(x)\, f_j(x)\, d\xi(x)$ is the information matrix of the experiment.

A fairly general problem in design theory is to minimize a convex function $\Psi(M)$ of the information matrix M. For example $\Psi(M) = \operatorname{tr} BM^{-1}$ for B positive semi-definite or $\Psi(M) = -\log |M|$ where $|M|$ denotes the determinant of M. Recently a number of equivalence theorems and closely related iterative procedures have appeared for minimizing $\Psi(M(\xi))$, see Kiefer [1973] for references. The purpose of this paper is to describe and study some very special iterative procedures which in approximation theory are called Remez type procedures or Remez exchange

---

*This research was supported by the National Science Foundation Grant 33552x2 and the Davis Ross Grant 6633. Reproduction is permitted in whole or in part for any purposes of the United States Government.

procedures. These procedure will be used to minimize $c'M^{-1}(\xi)\,c$ for a fixed vector $c' = (c_0, c_1, \ldots, c_k)$. In Section 2 we outline and discuss the procedure and give two simple examples. Section 3 contains a proof of the convergence. This proof as well as the procedure is taken from Meinardus [1967]. The proof is given in a design theory context and is included here for completeness. Some geometrical aspects of the procedure are included in Section 4.

2. Remez Procedure.

One of the general iterative procedures for minimizing $\Psi(M)$ 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". That is, we choose $\bar{M}_n$ so that $g(\alpha) = \Psi((1-\alpha)M_n + \alpha\bar{M}_n)$ has a minimum derivative at $\alpha = 0$. Then $M_{n+1} = (1-\alpha)M_n + \alpha\bar{M}_n$ and α is suitably chosen to give a decrease in Ψ. Since the set of all information matrices $M(\xi)$ is "spanned" by the set $M(\xi_x) = f(x)\,f'(x)$, $x \in \mathcal{X}$ (ξ_x cencentrates 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 then find the x value to give the minimum value for $g'(0)$ as a function of x. This result gives $g'(0) = \mathrm{tr}\,\nabla\Psi(M_n)(f(x)\,f'(x) - M_n)$ where $\nabla\Psi(M)$ is the kxk matrix with entries $(\nabla\Psi(M))_{ij} = \frac{\partial}{\partial m_{ij}}\Psi(M)$. We thus move in a direction $f(x)\,f'(x)$ where x minimizes $f'(x)\nabla\Psi(M_n)\,f(x)$. In certain special cases for Ψ the $\alpha = \alpha_n$ at the nth stage may be explicitly chosen in an optimal manner. The most general method is simply to use $\alpha_n \to 0$, to obtain some sort of convergence and $\Sigma\,\alpha_n = \infty$ to prevent convergence before reaching a minimum. In the case $\Psi(M) = c'M^{-1}c$ for some $c' = (c_0, c_1, \ldots, c_n)$ we obtain

$$(2.1) \qquad f'(x)\nabla\Psi(M)\,f(x) = -(f'(x)\,M^{-1}\,c)^2$$

so that $x = x_n$ is chosen to maximize $|f'(x)\,M_n^{-1}c|$. The Fedorov type procedure (see Fedorov [1972]), then chooses $\xi_{n+1} = (1-\alpha_n)\,\xi_n + \alpha_n\,\xi_{x_n}$, thus moving slowly toward a measure ξ concentrating mass on the extreme of $|f'(x)M_n^{-1}c|$. Part of the general equivalence theorem states that for ξ^* minimizing $\Psi(M(\xi))$, the value $g'(0)$ must be zero for all x. In the case $\Psi(M) = c'M^{-1}c$ this reduces to

$$(2.2) \qquad \left(c'M^{-1}(\xi^*)\ f(x)\right)^2 \le c'M^{-1}(\xi^*)c.$$

A presumably faster method, see Silvey and Titterington [1973], is to choose ξ_{n+1} to minimize $\Psi(M(\xi))$ where ξ is restricted to have support on the support of ξ_n plus the point x_n. Thus, if one starts with a measure ξ_0 with mass $p_1^{(0)}, \ldots, p_r^{(0)}$ on $x_1, \ldots x_r$, the point x_{r+1} is found maximizing $|f'(x)M^{-1}(\xi_0)\ c|$ and the values $p_i^{(1)}$ $i = 1,\ldots,r+1$ are found to minimize

$$(2.3) \qquad \Psi\left(\sum_{i=1}^{r+1} p_i^{(1)}\ f(x_i)\ f'(x_i)\right)$$

As one proceeds, most of the p_i values will be zero so that the number of effective x points remains bounded. In general this bound is $(k+1)(k+2)/2$. In the case $\Psi(M) = c'M^{-1}c$ an optimal design can always be found on k+1 points. This is due to a theorem of Elfving (see Karlin and Studden [1966]) which states that ξ^* minimizes $c'M^{-1}(\xi)c$ if and only if there exists a function ϵ with $|\epsilon(x)| = 1$ such that $\int \epsilon(x)\ f(x)\ d\xi(x) = \beta_* c$ for $\beta_*^{-2} = \min_\xi c'M^{-1}(\xi)c$ and $\beta_* c$ is a boundary point of a certain set R which is the convex hull of the set $\{\pm f(x) \mid x \in \mathcal{X}\}$

From the relation $\beta_*^{-2} = \inf_\xi c'M^{-1}(\xi)c$ it follows that

$$(2.4) \qquad \begin{aligned} \beta_*^{-2} &= \inf_\xi \sup_d \frac{(c'd)^2}{d'M(\xi)d} \\ &\ge \inf_\xi \frac{(c'd)^2}{\int (d'f(x))^2 d\xi(x)} \\ &= \frac{(c'd)^2}{\sup_x (d'f(x))^2} \end{aligned}$$

In this case

$$\text{(2.5)} \qquad \sup_x |d'f(x)| \geq \beta_*$$

for any d such that $c'd = 1$. Equation (2.4) provides the connection between the design theory and approximation theory since the sup and inf can be interchanged to show that the inf over d in (2.5) is β_*.

The Remez procedure for $\mathcal{X} = [a,b]$ restricts attention to ξ_n with support on k + 1 points and describes a method of "exchange".

One starts with a set $a \leq x_0^{(0)} < x_1^{(0)} <\ldots< x_k^{(0)} \leq b$. By the Elfving Theorem mentioned above applied to the set $\mathcal{X} = \{ x_\nu^{(0)}; \nu=0,\ldots,k\}$ the optimal weights $p_\nu^{(0)}$ $\nu = 0, 1,\ldots,k$ are a solution of

$$\text{(2.6)} \qquad \sum_{\nu=0}^{k} \epsilon_\nu^{(0)} p_\nu^{(0)} f(x_\nu^{(0)}) = \beta_0 c$$

where $\quad \epsilon_\nu^{(0)} = \pm 1, \; p_\nu^{(0)} \geq 0, \; \Sigma p_\nu^{(0)} = 1, \; \beta_0 > 0$

and β_0^{-2} is the minimum value of $c'M^{-1}(\xi)c$ for ξ with support on $x_i^{(0)}$, $i = 0,\ldots,k$. In general one must take the solution of (2.6) with the maximal β_0; if the $f(x_\nu^{(0)})$ are linearly independent the solution is then unique. Letting ξ_0 denote the above design and

$$\text{(2.7)} \qquad \varphi_0(x) = c'M^{-1}(\xi_0)\ f(x) \ / \ c'M^{-1}(\xi_0)c$$

(see (2.2)) we then choose a new set of points $a \leq x_0^{(1)} < x_1^{(1)} <\ldots< x_k^{(1)} \leq b$ so that

$$\text{(2.8)} \qquad \text{(1)} \quad |\varphi_0(x_\nu^{(1)})| \geq \beta_0, \; \nu = 0, 1,\ldots,k$$

$$\text{(2)} \quad |\varphi_0(x_{\nu_0}^{(1)})| > \beta_0 \text{ for some } \nu_0$$

$$\text{(3)} \quad \operatorname{sgn} \varphi_0(x_\nu^{(1)}) = \alpha \operatorname{sgn} \varphi_0(x_\nu^{(0)})$$

where α is constant $= \pm 1$.

The next design ξ_1 is then chosen by taking $p_\nu^{(1)}$ as a solution of

$$\sum_{\nu=0}^{k} \epsilon_\nu^{(1)} \; p_\nu^{(1)} \; f\,(x_\nu^{(1)}) = \beta_1 c$$

Continuing in this manner we obtain a sequence of designs ξ_n and values $\beta_n^{-2} = c'M^{-1}(\xi_n)c$ which hopefully converge.

With regard to the conditions (1) (2) and (3) for the new set of points there are two usual methods of proceding. Typically the function $\varphi_0(x)$ will have k-1 local extrema $x_i^{(1)}$, $i = 1,\ldots,k-1$ and one uses these together with $x_0^{(1)} = a$ and $x_k^{(1)} = b$. The other method is to just choose ξ to give $|\varphi_0(\xi)| = \max_x |\varphi_0(x)|$ and then exchange ξ for one of the $x_\nu^{(0)}$ values to satisfy (3) Roughly speaking this entails replacing ξ with an adjacent $x_\nu^{(0)}$ value for which φ_0 has the same sign. In general we use the following rule.

ξ value	sgn $\varphi_0(\xi)$ =	ξ replaces
$a \le \xi < x_0^{(0)}$	sgn $\varphi_0(x_0^{(0)})$	$x_0^{(0)}$
$a \le \xi < x_0^{(0)}$	-sgn $\varphi_0(x_0^{0})$	$x_k^{(0)}$
$0 \le \nu \le k-1$		
$x_\nu^{(0)} < \xi < x_{\nu+1}^{(0)}$	sgn $\varphi_0(x_\nu^{(0)})$	$x_\nu^{(0)}$
$x_\nu^{(0)} < \xi < x_{\nu+1}^{(0)}$	-sgn $\varphi_0(x_\nu^{(0)})$	$x_{\nu+1}^{(0)}$
$x_k^{(0)} < \xi \le b$	sgn $\varphi_0(x_k^{(0)})$	$x_k^{(0)}$
$x_k^{(0)} < \xi \le b$	-sgn $\varphi_0(x_k^{(0)})$	$x_0^{(0)}$

Note that in both of these casesone of the $x_\nu^{(0)}$ values is replaced by the ξ value for which $|\varphi_0(\xi)| = ||\varphi_0|| = \sup_x |\varphi_0(x)|$. Something of this nature is necessary in order to prevent convergence before reaching the required limit.

We will prove convergence of the above procedure for the case where the vector c is "Tchebycheffian" with respect to the system $f_i(x)$, $i = 0, 1,\ldots,k$. This means that for every set of k + 1 points $a \le x_0 < x_1 <\ldots< x_k \le b$ the determinants $D_\nu(c) = D_\nu(c; x_0, x_1,\ldots,x_k) = |f(x_0), f(x_1),\ldots,f(x_{\nu-1}), c, f(x_{\nu+1}),\ldots,f(x_k)|$ are never zero and they **alternate in sign**. We now show under these conditions and (1) (2) and (3) that $\beta_{n+1} \ge \beta_n$ or $c'M^{-1}(\xi_{n+1})c \le c'M^{-1}(\xi_n)c$. As inspection of the equations (2.6) shows that the values $\epsilon_\nu^{(0)}$ $\nu = 0, 1,\ldots,k$ alternate in sign. Moreover by (2.6) and (2.7) $\sum_\nu p_\nu^{(0)} \epsilon_\nu^{(0)} \varphi_0(x_\nu^{(0)}) = \beta_0$ and by (2.2), $|\varphi_0(x_\nu^{(0)})| \le \beta_0$ so that $\epsilon_\nu^{(0)} \varphi_0(x_\nu^{(0)}) = \beta_0$. This implies that

(2.9) $\varphi_0(x_\nu^{(0)})$ $\nu = 0,1,\ldots,k$

alternate in sign

These above conclusions hold at each step so that

$$\beta_1 = \sum_\nu \epsilon_\nu^{(1)} p_\nu^{(1)} \varphi_0(x^{(1)})$$

$$= \sum_\nu \epsilon_\nu^{(1)} p_\nu^{(1)} |\varphi_0(x^{(1)})| \operatorname{sgn} \varphi_0(x_\nu^{(1)})$$

$$= \sum_\nu \epsilon_\nu^{(1)} p_\nu^{(1)} |\varphi_0(x^{(1)})| \operatorname{sgn} \varphi_0(x_\nu^{(0)})$$

$$= \sum_\nu p_\nu^{(1)} |\varphi_0(x_\nu^{(1)})|$$

Therefore

(2.10) $$\beta_1 = \beta_0 + \Sigma\, p_\nu^{(1)} \{|\varphi_0(x_\nu^{(1)})| - \beta_0\}$$

By condition (1) that $|\varphi_0(x_\nu^{(1)})| \ge \beta_0$ we have

(2.11) $$\beta_1 \ge \beta_0.$$

We should note here that the Silvey and Titterington type procedure would choose the "best" subset of k + 1 points from $\{x_0^{(0)} \ldots x_{k+1}^{(0)}, \xi\}$ whereas the Remez procedure is not generally the best but the exchange is made explicit. Thus instead of determining the β in (2.6) for each subset of k + 1 points an exchange is made and the system of equations (2.6) is solved once instead of n + 1 times. The sacrifice is, of course, a smaller increase in the β value.

Example 1. This example will be used to illustrate the choice of exchange points. Let $\mathcal{X} = [-1, 1]$, $f'(x) = (1,x)$ and $c = (0,1)$. For an initial two points we use $x_0^{(0)} = -1/2$ and $x_1^{(0)} = +3/4$: Then $\varphi_0(x) = x - 1/8$ and $\xi = -1$ giving $|\varphi_0(\xi)| = \max_x |\varphi_0(x)|$. Moreover $\epsilon_0^{(0)} = -1 = \text{sgn}\, \varphi_0(x_0^{(0)})$, $\epsilon_1^{(0)} = +1 = \text{sgn}\, \varphi_0(x_1^{(0)})$ and $\beta_0 = 5/8$. One can easily show that $\xi = -1$ must be exchanged with $x_0^{(0)} = -1/2$ giving $\beta_1 = 7/8$. The exchange with $x_1^{(0)}$ gives a decrease to $\beta_1 = 1/4$. The next step will produce $\xi = +1$. One could exchange $x_0^{(0)}$ and $x_1^{(0)}$ at the first step for the two extreme of $|\varphi_0(x)|$, namely $x = \pm 1$.

Example 2. Let $f'(x) = (1, x, x^2, (x - \eta)_+^2)$ for $\mathcal{X} = [-1,1]$, where $(x - \eta)_+^2 = (x - \eta)^2$ if $x \geq \eta$ and equals zero for $x < \eta$. We consider the case $\eta = 0.4$. The procedure is terminated if the critical value

$$\epsilon_n = \frac{||\varphi_n|| - \beta_n}{\beta_n} \leq \epsilon = 10^{-5}$$

where $||\varphi_n|| = \max_x |\varphi_n(x)|$. Four equally spaced points on $[-1,1]$ where used for an initial set x_ν^0, $\nu = 0,1,2,3$. The results are as follows.

n	$x_0^{(n)}$	$x_1^{(n)}$	$x_2^{(n)}$	$x_3^{(n)}$	β_n	ϵ_n
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}

The design ξ_3 is then

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

and $\beta_n^{-2} = 247.7$

The Fedorov procedure for this example was run for 30 iterations and "rounded off" to a four point design as described in Fedorov [1972] page 109. The results produced a design

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

and $c'M^{-1}(\tilde{\xi}_{30})c = 267.9$

It should be remarked that each iteration in the Fedorov procedure usually takes less time than an iteration using the Remez procedure.

§ 3 Proof of Convergence. We assume that the conditions (2.8) hold, that c is Tchebycheffian with respect to $\{f_i\}$, and that ξ giving $\max_x |\varphi(x)|$ is one of the points in the exchange.

We take equation (2.10) with 0 and 1 replaced by n and n + 1 to give

$$\beta_{n+1} - \beta_n = \sum_\nu p_\nu^{(n+1)} \{|\varphi_n(x_\nu^{(n+1)})| - \beta_n\}$$

This implies $\beta_{n+1} \geq \beta_n$. Since at each stage there exists a ν_n such that $\varphi_n(x_{\nu_n}^{(n+1)}) = ||\varphi_n|| = \sup_x |\varphi_n(x)|$ it then follows that

$$\beta_{n+1} - \beta_n \geq p_{\nu_n}^{(n+1)} \{||\varphi_n|| - \beta_n\} \tag{3.1}$$

We will show subsequently that $\underline{\lim}_n \; p_\nu^{(n)} > 0$ for each ν. Since the β_n are bounded by β_* they must converge and hence $||\varphi_n|| - \beta_n \to 0$. By the definition of φ_n given in (2.7) by

$$\varphi_n(x) = c'M^{-1}(\xi_n)\ f(x)\ /\ c'M^{-1}(\xi_n)c$$

it follow from (2.5) that

$$\beta_* \leq ||\varphi_n||$$

An upper bound on $||\varphi_n||$ can be obtained from equation (3.1) to give

$$\beta_* \leq ||\varphi_n|| \leq \beta_n + (\beta_{n+1} - \beta_n) \ / \ p_{\nu_n}^{(n+1)}$$
$$\leq \beta_* + (\beta_{n+1} - \beta_n) \ / \ p_{\nu_n}^{(n+1)}$$

Therefore $||\varphi_n||$ and hence β_n converges to β_*.

In order to show that $\underline{\lim}_n \, p_\nu^{(n)} > 0$ we first show that $\underline{\lim}_n \, |x_{\nu+1}^{(n)} - x_\nu^{(n)}| > 0$. In the contrary case there exists a ν_0 and a subsequence such that $x_{\nu_0+1}^{(n)} - x_{\nu_0}^{(n)} \to 0$ along the subsequence. We further refine the subsequence so that all $x_\nu^{(n)}$ converge. The limit set will have at most k points say $z_1, z_2, \ldots, z_k$. We then choose a polynomial $a'f(x)$ such that

$$a'f(z_\nu) = 0 \quad \nu = 1, 2, \ldots, k$$
$$a'c = 1$$

then from the equation

$$\text{(3.2)} \qquad \sum_\nu p_\nu^{(n)} \epsilon_\nu \, f(x_\nu^{(n)}) = \beta_n c$$

we obtain

$$\sum_\nu p_\nu^{(n)} \epsilon_\nu \, a'f(x_\nu^{(n)}) = \beta_n$$

However, in this case, the left side goes to zero from the continuity of the functions f_i while the right side β_n increases to $\beta_* > 0$. The resulting contradiction gives $\underline{\lim}_n \, x_{\nu+1}^{(n)} - x_\nu^{(n)} > 0$. Now from each ν and n we choose the vector a_n so that

$$\text{(3.3)} \qquad \begin{aligned} a_n' \ f(x_i^{(n)}) &= 0 \qquad\qquad i = 0, 1, \ldots, k, \ i \neq \nu \\ a_n' \ c &= 1 \end{aligned}$$

Then $\quad p_\nu^{(n)} \epsilon_\nu^{(n)} a_n' \, f(x_\nu^{(n)}) = \beta_n$. If $\underline{\lim}_n \, p_\nu^{(n)} = 0$ then on a suitable subsequence

$a_n' f(x_\nu^{(n)}) \to \infty$. However the solution a_n from (3.3) is easily seen to be bounded if $\underline{\lim}_n |x_{\nu+1}^{(n)} - x_\nu^{(n)}| \geq \eta > 0$.

§ 4 Geometry of Remez Procedure. A number of interpretations are available here. As remarked around (2.4) and (2.5) the design problem is equivalent to minimizing $\sup_x |d'f(x)|$ subject to $d'c = 1$. For $X_0 = \{x_0^{(0)}, x_1^{(0)},\ldots,x_k^{(0)}\}$, $d' = c'M^{-1}(\xi_0) / c'M^{-1}(\xi_0)c$ gives minimum value for $\sup_{x\in X_0} |d'f(x)|$. One then "plots" $\varphi_0(x)$ or considers $|\varphi_0(x)|$ to find it's maximum ξ. An exchange is then made to give X_1 etc. One is actually solving for the $k + 1$ points such that $\inf_d \sup_{x\in X_i} |d'f(x)|$ is a maximum. This turns out to be equivalent to finding $\inf_d \sup_x |d'f(x)|$. The Remez procedure can be readily interpreted using the Elfving Theorem. The vector $d' = c'M^{-1}(\xi_0) / c'M^{-1}(\xi_0)c$ gives a support plane $d'z = \beta_0$ $(z = (z_0, z_1,\ldots,z_k))$ at $\beta_0 c$ to the set $R^{(0)}$ determined as the convex hull of $\pm f(x_i^{(0)})$, $i = 0, 1,\ldots,k$. The representation

$$\beta_0 c = \sum_\nu p_\nu \epsilon_\nu f(x_\nu^{(0)})$$

gives $\beta_0 c$ as a convex combination of $\epsilon_\nu f(x_\nu^{(0)})$ and such $\epsilon_\nu f(x_\nu^{(0)})$ lies in the hyperplane $d'z = \beta_0$, i.e. $\epsilon_\nu \varphi_0(x_\nu^{(0)}) = \beta_0$. One now chooses ξ giving maximum value for $|\varphi_0(x)|$ so that $\epsilon\varphi_0(\xi) > \beta_0$ or $\epsilon f(\xi)$ lies on the side of hyperplane $d'z = \beta_0$ opposite the origin. If one can now exchange $\epsilon f(\xi)$ with one of the vectors $\epsilon_i f(x_i^{(0)})$ so $\beta_1 c$ $(\beta_1 > 0)$ is a convex combination of the new set of vectors then "clearly" $\beta_1 \geq \beta_0$. This is true since if $d'c = 1$ then

$$\begin{aligned}\beta_1 = \beta_1 d'c &= d'\beta_1 c \\ &= d'\left(\sum_{\nu\neq i} p_\nu^{(1)} \epsilon_\nu f(x_\nu^{(0)}) + p_i^{(1)} \epsilon f(\xi)\right) \\ &= d'\left(\sum_\nu p_\nu^{(1)} \epsilon_\nu f(x_\nu^{(0)}) + p_i^{(1)} (\epsilon f(\xi) - \epsilon_i f(x_i^{(0)}))\right)\end{aligned}$$

$$= \sum_{\nu} p_{\nu}^{(1)} \epsilon_{\nu} \varphi_0(x_{\nu}^{(0)}) + p_i^{(1)} (\epsilon\varphi(\xi) - \beta_0)$$

$$= \beta_0 + p_i^{(1)} (|\varphi(\xi)| - \beta_0)$$

If one exchanges more than one point we end up with equation (2.10).

In order to determine how the exchange should be made we let $a_{\nu} = \epsilon_{\nu} f(x_{\nu}^{(0)})$ and $a = \epsilon\, f(\xi)$. Then

(4.1) $$\beta_0 c = \sum_{\nu=0}^{k} p_{\nu}\, a_{\nu} \qquad (p_{\nu} = p_{\nu}^{(0)})$$

and we wish an exchange so that a similar equation holds. One simply takes a representation

(4.2) $$a = \sum_{\nu} q_{\nu}\, a_{\nu}$$

and considers an exchange using any a_i with $q_i \neq 0$. Solving (4.2) for a_i, and substituting in (4.1) gives

$$\beta_0 c = \sum_{\nu \neq i} p_{\nu}^{(0)} a_{\nu} + p_i (a - \sum_{\nu \neq i} q_{\nu}\, a_{\nu}) / q_i$$

$$= \sum_{\nu \neq i} q_{\nu} \left(\frac{p_{\nu}}{q_{\nu}} - \frac{p_i}{q_i} \right) a_{\nu} + \frac{p_i}{q_i}\, a$$

In order to have all the coefficients non-negative we choose i to give minimum value for p_i/q_i for $q_i > 0$. A renormalization then produces

$$\beta_1 c = \sum_{\nu \neq i} p_{\nu}^{(1)} a_{\nu} + p_i^{(1)} a$$

This method of exchange has certain advantages over the one indicated in the table in § 2. One advantage is that the ordering of the x values is not used so that we do not require $\mathcal{X}$ to be an interval.

REFERENCES

Fedorov, V.V. (1972) Theory of Optimal Experiments, Academic Press, New York.

Karlin, S. and Studden, W. J., (1966), Optimal Experimental Designs, Ann. Math. Statist., 34, 783-815.

Kiefer, J., (1973), General Equivalence Theory for Optimum Designs (Approximate Theory) (To appear).

Meinardus, G., (1967), Approximation of Functions; Springer-Verlag, New York.

Silvey, J. D. and Titterington, D. M. (1973), A Geometric Approach to Optimal Design Theory, Biometrica 60, 21 - 32.