The Annals of Statistics
1980, Vol. 8, No. 3, 652–663

OPTIMAL DESIGNS FOR SECOND ORDER PROCESSES WITH GENERAL LINEAR MEANS

BY CARL SPRUILL[1]

Georgia Institute of Technology

For each x in some factor space X an experiment can be performed whose outcome is $\{Y(x, t) : t \in T]$ where $Y(x, t) = m_x(\theta, t) + \varepsilon(t)$. The zero mean error process $\varepsilon(t)$ has known covariance function K and the maps m_x (of known form) are linear from the parameter space Θ to the rkhs generated by K. Expressions for the variance of the umvlue of $\tau(\theta)$ (where τ is linear) are given which are analogous to the formulas in the finite dimensional Θ case. An Elfving's theorem is proved and a number of examples are given.

1. Introduction. The action of an unknown variable force $\{\theta(t) : t \in [0, 1]\}$ on a particle of known mass $x \in [a, b]$ which is initially at rest may be observed repeatedly. The results are marred by an observational error of zero mean and known covariance. In an experiment with N uncorrelated observations $\{Y(x_i, t) : t \in [0, 1]\}_{i=1}^N$ of the position function of the particle over time, what is the "best" selection of masses $\{x_1, \cdots, x_N\}$ for estimating $\int_0^1 \theta(s)\, ds$ (see Example 2.1)?

The problem described above is a particular case of a type of design problem whose solution is characterized below. In the case of the general problem, for each x in a set X of possible levels of feasible experiments, an experiment can be performed whose outcome is a stochastic process $\{Y(x, t) : t \in T\}$. It is assumed that the process has a mean function $m(x, \theta, t)$ of known form, linear in the unknown parameter θ. The parameter θ is an element of a linear, but otherwise arbitrary, space Θ. For each x in X and θ in Θ the function $m(x, \theta)$ on T is a member of the reproducing kernel Hilbert space $H(K)$ generated by the known covariance kernel $K(s, t) = \text{Cov}[Y(x, s), Y(x, t)]$, $x \in X$, $s, t \in T$. The value $\tau(\theta)$, where τ is a linear (not necessarily continuous) functional on Θ, is to be estimated on the basis of N uncorrelated observations $\{Y(x_i, t) : t \in T, i = 1, \cdots, N\}$. The problem to be solved is to find the experimental design which minimizes the variance of the best linear unbiased estimator of $\tau(\theta)$.

The following publications and references therein contain information on related optimum design problems: Fedorov (1972), Kiefer (1974), Mehra (1974), Pazman (1978), Spruill and Studden (1978), and Wahba (1976).

The model given above generalizes the model considered in Spruill and Studden (1978) by allowing a more general mean and parameter space. For the most part, the development below parallels that in Spruill and Studden. There is at least one major difference however. Under assumptions similar to those in the finite-dimen-

Received April 1978; revised September 1978.

[1]Research support was provided by NSF Grant No. MCS 76-11040-A01.

AMS 1970 *subject classifications.* Primary, 62J05, 62K05; secondary 62M99.

Key words and phrases. Linear operator, linear space, mvlue, optimum designs, kernel Hilbert space.

sional case, the variance of the blue may fail to be minimized by any probability measure on the factor space which concentrates on a finite number of points. A design is called optimal below only if it minimizes the variance of the blue and is supported on a finite number of points.

2. Preliminary results. Consider the discrete design ξ which places masses $p_i = \frac{n_i}{N}$ at x_i, $i = 1, \cdots, r$, where $x_i \in X$ and $\{n_i\}_{i=1}^r$ are integers with $\Sigma_{i=1}^r p_i = 1$. The experiment consists of taking N uncorrelated observations of the process, n_i at x_i. Set $\Gamma = \{(x_1, 1), (x_1, 2), \cdots, (x_r, n_r)\} \times T$ and define the process $\{Z(\gamma) : \gamma \in \Gamma\}$ by $Z(\gamma) = Y_j(x_i, t)$, the jth observation at x_i, where $\gamma = ((x_i, j), t)$. The covariance kernel of Z is given by

$$\begin{aligned} B(\gamma_1, \gamma_2) &= K(t_1, t_2) \qquad (x_{i_1}, j_1) = (x_{i_2}, j_2) \\ &= 0 \qquad\qquad\quad \text{otherwise} \end{aligned}$$

where $\gamma = ((x_{i_l}, j_l), t_l)$, $l = 1, 2$. Denote by $\langle Z, g\rangle_B$ the random variable in $L_2[Z(\gamma) : \gamma \in \Gamma]$ which is the image of g in the reproducing kernel Hilbert space $H(B)$ associated with B. The class of linear estimators of $\tau(\theta)$ is $\{\langle Z, g\rangle_B : g \in H(B)\}$. The functional $\tau(\theta)$ is said to be (linearly) estimable with respect to this design if there is a $g \in H(B)$ such that $E_\theta\langle Z, g\rangle_B \equiv \tau(\theta)$. Denote by $(\cdot\ ,\ \cdot)_B$ the inner product on $H(B)$ and for ξ as given above fixed, by m the map from Θ to $H(B)$ defined by $m(\theta)(\gamma) = m(x_i, \theta, t)$ if $\gamma = ((x_i, j), t)$. Parzen (1959) proved the following.

THEOREM (10A). *Given that $\tau(\theta)$ is estimable, there is a unique linear estimator $\langle Z, g_0\rangle_B$ which is the uniformly minimum variance linear unbiased estimator* (umvlue) *of $\tau(\theta)$ with variance $\| g_0\|_B^2$. Furthermore $\langle Z, g\rangle_B$ is the* umvlue *of $\tau(\theta)$ if and only if g is the unique function in the closure of $\mathfrak{R}(m) = \{m(\theta) : \theta \in \Theta\}$ satisfying $\tau(\theta) \equiv (m(\theta), g)_B$, $\theta \in \Theta$.*

Using the methods of Parzen (1959), one can prove the following.

LEMMA 2.1. *The element g is in $H(B)$ if and only if $g((x_i, j), \cdot) \in H(K)$ for (x_i, j), $i = 1, \cdots, r$, $j = 1, \cdots, n_i$. Furthermore, if g and h are in $H(B)$*

$$(g, h)_B = \Sigma_{i=1}^r \Sigma_{j=1}^{n_i} (g(x_i, j), h(x_i, j))_K.$$

Let m' be the map from $H(B)$ to the space of linear functionals Θ' on Θ defined by

$$m'(v)(\theta) = (v, m(\theta))_B. \tag{1}$$

The map m' is the restriction of the usual transpose to $H(K)$ (see Taylor (1958)). Setting

$$M = N^{-1} m'm \tag{2}$$

it will be shown below that an expression for the variance of the umvlue of $\tau(\theta)$ may be obtained which is analogous to that obtained for $\Theta = R^{k+1}$ in Spruill and

Studden (1978). If $\Theta = R^{k+1}$ the variance of the umvlue of $\tau(\theta)$ is $N^{-1}\tau' M^{+}\tau$ where M^{+} is the Moore-Penrose generalized inverse of M.

The notion of a generalized inverse extends to mappings between arbitrary linear spaces (see Nashed and Votruba (1976)). In particular, we use the notion of an algebraic generalized inverse (A.G.I.). The reader is referred to Proposition 1.16 and the material preceding Proposition 1.17 of Nashed and Votruba (1976) which shows that given the linear operator $L : \Theta \to \Theta'$ there are (algebraic) projectors P and Q, P defined on Θ and Q on Θ', such that

$$\begin{aligned} \Theta &= \mathcal{N}(L) \dotplus \mathcal{M}, \\ \Theta' &= \mathcal{R}(L) \dotplus \mathcal{S}, \\ \mathcal{R}(P) &= \mathcal{N}(L), \ \mathcal{R}(I-P) = \mathcal{M}, \\ \mathcal{R}(Q) &= \mathcal{R}(L), \quad \text{and} \quad \mathcal{R}(I-Q) = \mathcal{S}. \end{aligned} \tag{3}$$

The symbol $\dotplus$ means algebraic direct sum so that P and Q are not necessarily continuous projectors. Indeed, we should emphasize that no topological assumptions have been made, or will be made in this section. For a mapping A, $\mathcal{R}(A)$ denotes the range of A and $\mathcal{N}(A)$ denotes the null space of A. Nashed and Votruba demonstrate the existence of a unique (for P and Q fixed) linear operator $L^{\#} : \Theta' \to \Theta$ called the A.G.I. of L which satisfies

$$\begin{aligned} LL^{\#}L &= L \\ L^{\#}LL^{\#} &= L^{\#} \\ L^{\#}L &= I - P \\ LL^{\#} &= Q. \end{aligned} \tag{4}$$

We note that if $\theta' = L\theta$ then $\theta - L^{\#}\theta' = (I - L^{\#}L)\theta$ or

$$\theta = L^{\#}\theta' + (I - L^{\#}L)\theta. \tag{5}$$

If the form $\tau(\theta)$, $\tau \in \Theta'$, is estimable then there is a umvlue which corresponds to, say $g_0 \in \overline{\mathcal{R}(m)}$ (the closure of $\mathcal{R}(m)$). Let $\theta_n \in \Theta$ be such that $\lim_{n\to\infty}\|m(\theta_n) - g_0\|_B = 0$ and define $\theta_n' = L\theta_n$, where $L = NM$.

THEOREM 2.1. (i) *The form $\tau(\theta)$ is estimable if and only if $\tau \in \mathcal{R}(m')$.*
(ii) *If $\tau(\theta)$ is estimable then $\|g_0\|_B^2 = N^{-1}\lim_{n\to\infty}\theta_n'(M^{\#}\theta_n')$.*

PROOF. (i) $\tau(\theta) \equiv (m(\theta), g_0)_B$ if and only if $\tau(\theta) \equiv m'g_0(\theta)$.

(ii)
$$\begin{aligned} \|g_0\|_B^2 &= \lim_{n\to\infty}\|m(\theta_n)\|_B^2 \\ &= \lim_{n\to\infty} m'm(\theta_n)(\theta_n) \\ &= \lim_{n\to\infty} L(\theta_n)(\theta_n) \\ &= \lim_{n\to\infty} \theta_n'(L^{\#}\theta_n' + (I - L^{\#}L)\theta_n) \\ &= \lim_{n\to\infty} \theta_n'(L^{\#}\theta_n') + \theta_n'(I - L^{\#}L)\theta_n. \end{aligned}$$

The result will follow as soon as it is shown that $\theta_n'(I - L^{\#}L)\theta_n = 0$. To see this, suppose $u \in \mathfrak{N}(L)$. Then for all $y \in \Theta$, $0 = Lu(y) = N(m(u), m(y))_B$. In particular, $m(u) = 0$ so $\mathfrak{N}(L) \subset \mathfrak{N}(m)$. Since $(I - L^{\#}L)\theta_n \in \mathfrak{N}(L)$ and

$$\theta_n'(I - L^{\#}L)\theta_n = N\big(m(\theta_n), m(I - L^{\#}L)\theta_n\big)_B$$

the lemma has been proved.

COROLLARY. *If $\tau(\theta)$ is estimable and $\mathfrak{R}(m)$ is closed in $H(B)$ the variance of the* umvlue *of $\tau(\theta)$ is given by*

$$V = N^{-1}\tau(M^{\#}\tau).$$

Denote by Ξ the set of all probability measures on X with finite support. Let H_ξ be the Hilbert space of real valued functions f on $S(\xi)xT$, where $S(\xi)$ is the support of ξ, such that $f(x_i, \cdot) \in H(K)$ for $x_i \in S(\xi)$. The inner product of f and g is

$$(f, g)_\xi = \Sigma_{x \in S(\xi)}\xi(x)(f(x), g(x))_K.$$

Denote by m_ξ the map from Θ to H_ξ defined by $m_\xi(\theta)(x_i, t) = m(x_i, \theta, t)$ for $x_i \in S(\xi)$, $t \in T$. Let

$$M(\xi) = m_\xi' m_\xi. \tag{6}$$

If ξ has all its mass at a point $x \in X$ write m_x rather than m_ξ.

When ξ has rational probabilities at its support points, (2) and (6) coincide. We are now in a position to make the following definitions for $\xi \in \Xi$.

DEFINITION. The linear form $\tau(\theta)$ is estimable with respect to the design $\xi \in \Xi$ if $\tau \in \mathfrak{R}(m_\xi')$.

Let $M_\xi = \{m_\xi(\theta) : \theta \in \Theta\}$. Then $M_\xi \subset H_\xi$ and if $\tau(\theta)$ is estimable there is a unique $g \in \bar{M}_\xi$ (the closure of M_ξ) such that $(m_\xi(\theta), g)_\xi \equiv \tau(\theta)$ and a sequence $\{\theta_n\} \subset \Theta$ such that $\|m_\xi(\theta_n) - g\|_\xi \to 0$ as $n \to \infty$. Let $\theta_n' = M(\xi)\theta_n$ and define

$$d(\tau, \theta) = \lim_{n\to\infty}\theta_n' M^{\#}(\xi)\theta_n'. \tag{7}$$

If τ is not estimable let $d(\tau, \xi) = +\infty$.

DEFINITION. The design $\xi_0 \in \Xi$ is said to be optimal for estimating $\tau(\theta)$ if $d(\tau, \xi_0) = \inf_\Xi d(\tau, \xi)$.

LEMMA 2.3. *For $\xi \in \Xi$ the operator $M(\xi)$ is given by $M(\xi) = \int m_x' m_x \, d\xi(x)$.*

PROOF. For u and v in Θ

$$\begin{aligned}(M(\xi)u)v &= (m_\xi' m_\xi u)v = (m_\xi(u), m_\xi(v))_\xi \\ &= \int (m(x, u), m(x, v))_K \, d\xi(x) \\ &= \int (m_x'(m_x u))v \, d\xi(x).\end{aligned}$$

LEMMA 2.4. *For any $\theta_1, \theta_2 \in \Theta$ and $\xi \in \Xi$*

$$|(M(\xi)\theta_1)\theta_2| \leqslant ((M(\xi)\theta_1)\theta_1)^{\frac{1}{2}}((M(\xi)\theta_2)\theta_2)^{\frac{1}{2}}$$

with equality if and only if $m_\xi(\theta_1) = k m_\xi(\theta_2)$ for some constant k.

PROOF. Set $\|m_\xi(\theta_1) - sm_\xi(\theta_2)\|_\xi^2 = f(s)$. Then the real valued function f defined on $(-\infty, +\infty)$ has a minimum at

$$s_0 = \frac{(M(\xi)\theta_2)\theta_1}{(M(\xi)\theta_2)\theta_2} \text{ if } m_\xi(\theta_2) \neq 0.$$

Using the fact that $f(s) \geqslant 0$, with equality if and only if $m_\xi(\theta_1) = km_\xi(\theta_2)$, the result is proved as in the usual finite dimensional case.

In addition to the algebraic generalized inverse of a linear operator T, it is useful to have another notion of the transpose. It differs somewhat from both the usual transpose (see Taylor (1958), Chapter 1) and the transpose defined above. In particular, we suppose that $T : \Theta \to \Theta'$, where as above, Θ is a linear space and Θ' is the space of all linear functionals on Θ.

DEFINITION. The t-transpose of T, written T^t, is the linear operator mapping Θ into Θ' defined by $Tx(y) = T^ty(x)$, for all x, y in Θ.

The following lemma is true for T^t but is not in general true for the usual transpose. Some additional notation is required. For an arbitrary subset $A \subset \Theta$ let $A^\perp = \{\theta' \subset \Theta' : \theta'(\theta) = 0 \text{ for all } \Theta \in A\}$. For an arbitrary subset $B \subset \Theta'$ let $B^\perp = \{\theta \in \Theta : \theta'(\theta) = 0 \text{ for all } \theta' \in B\}$.

LEMMA 2.5.
(i) $\mathfrak{R}^\perp(T) = \mathfrak{N}(T^t)$
(ii) $\mathfrak{R}^\perp(T^t) = \mathfrak{N}(T)$
(iii) $\mathfrak{R}(T^t) = \mathfrak{N}^\perp(T)$
(iv) $\mathfrak{R}(T) = \mathfrak{N}^\perp(T^t)$.

PROOF. Parts (i) and (ii) are proved in a similar manner, so we only prove (i). Let $x \in \mathfrak{R}^\perp(T)$. Then $(Ty)x = 0$ for all $y \in \Theta$. Thus $T^tx(y) = 0$ for all $y \in \Theta$. Therefore $T^tx = 0$. The reverse inclusion follows from the same argument.

If W is any subspace contained in Θ then $W = W^{\perp\perp}$ (see Taylor (1958), Theorem 1.9-A). Parts (iii) and (iv) follow immediately.

LEMMA 2.7. *For any* $\theta' \in \Theta'$, $\theta' \neq 0$, *and* $\xi \in \Xi$

$$d(\theta', \xi) = \sup_{\theta \in U} \frac{(\theta'(\theta))^2}{(M(\xi)\theta)\theta}, \tag{8}$$

where $U = \{\theta : (M(\xi)\theta)\theta \neq 0\}$.

PROOF. As in (3), write $\Theta = \mathfrak{N}(M) \dotplus \mathfrak{M}$. First, suppose $\theta' \in \mathfrak{R}(M)$. Then $\theta' = MM^{\#}\theta'$ since $MM^{\#}$ is the projection onto $\mathfrak{R}(M)$. Thus

$$\begin{aligned}(\theta'(\theta))^2 &= [(MM^{\#}\theta')\theta]^2 \\ &= (M(M^{\#}\theta')\theta)^2 \\ &\leqslant [M(M^{\#}\theta')M^{\#}\theta'][(M\theta)\theta] \\ &= [\theta'(M^{\#}\theta')][(M\theta)\theta],\end{aligned}$$

for any θ. If $\theta \in U$ then $(M\theta)\theta \neq 0$ so

$$\frac{(\theta'(\theta))^2}{(M\theta)\theta} \leqslant (M^{\#}\theta')\theta'. \tag{9}$$

Setting $\theta = M^{\#}\theta'$, which is in U, equality is achieved in (9), proving the result for $\theta' \in \mathfrak{R}(M)$.

If $\theta' \in \mathfrak{R}(m') - \mathfrak{R}(M)$ then $\theta'(\theta)$ is estimable with umvlue corresponding to some $g_0 \in \overline{M}_\xi \subset H(\xi)$. Let $\{\theta_n\} \subset \Theta$ be such that $\|m(\theta_n) - g_0\|_\xi \to 0$ and set $\theta_n' = M\theta_n$. The linear form $\theta_n'(\theta)$ is estimable and its umvlue corresponds to $m(\theta_n)$ since

$$\theta_n'(\theta) = (m(\theta_n), m(\theta))_\xi$$

and $m(\theta_n) \in M_\xi$. Furthermore $\theta_n' \in \mathfrak{R}(M)$, so by the above argument

$$\begin{aligned} &d(\theta_n', \xi) = \sup_{\theta \in U} \frac{(\theta_n'(\theta))^2}{(M\theta)\theta} \qquad \text{and} \\ &d(\theta_n', \xi) = \|m(\theta_n)\|^2 \to \|g_0\|^2 = d(\theta', \xi). \end{aligned} \tag{10}$$

Notice that for fixed θ

$$\begin{aligned} \left| \frac{[\theta_n'(\theta)]^2}{(M\theta)\theta} - \frac{[\theta'(\theta)]^2}{(M\theta)\theta} \right| &= \left| \frac{(m(\theta_n), m(\theta))_\xi^2 - (g_0, m(\theta))_\xi^2}{\|m(\theta)\|_\xi^2} \right| \\ &= \left| \frac{(m(\theta_n) - g_0, m(\theta))_\xi (m(\theta_n) + g_0, m(\theta))_\xi}{\|m(\theta)\|_\xi^2} \right| \\ &\leqslant \|m(\theta_n) - g_0\|_\xi \|m(\theta_n) + g_0\|_\xi, \end{aligned} \tag{11}$$

so that given $\varepsilon > 0$ it follows from (11) that for n sufficiently large $\left| d(\theta_n', \xi) - \sup_{\theta \in U} \frac{[\theta'(\theta)]^2}{(M\theta)\theta} \right| < \varepsilon/2$ and from (10) that $|d(\theta_n', \xi) - d(\theta', \xi)| < \varepsilon/2$. Hence for n sufficiently large

$$\left| d(\theta', \xi) - \sup_{\theta \in U} \frac{[\theta'(\theta)]^2}{(M\theta)\theta} \right| < \varepsilon.$$

This proves the lemma for $\theta' \in \mathfrak{R}(m')$.

If $\theta' \notin \mathfrak{R}(m')$ then $\theta' = \theta'_{\mathfrak{R}(m')} + \theta'_{\mathfrak{T}}$ where $\Theta' = \mathfrak{R}(m') \dotplus \mathfrak{T}$ and $\theta'_{\mathfrak{T}} \neq 0$. If $\theta'_{\mathfrak{T}}(\theta) \equiv 0$ for $\theta \in \mathfrak{N}(M)$ then $\theta'_{\mathfrak{T}} \in \mathfrak{N}^{\perp}(M)$. By lemma 2.5 $\theta'_{\mathfrak{T}} \in \mathfrak{R}(M^t)$. Since $M\theta_1(\theta_2) = (m(\theta_1), m(\theta_2))_\xi$ and $M^t\theta_1(\theta_2) = M\theta_2(\theta_1) = (m(\theta_1), m(\theta_2))_\xi$ we observe that $M = M^t$ and hence that $\mathfrak{R}(M^t) = \mathfrak{R}(M)$. Since $\mathfrak{R}(M) \subset \mathfrak{R}(m')$, we have arrived at the contradictory conclusion that $\theta'_{\mathfrak{T}} \in \mathfrak{R}(m')$ unless θ_0 in $\mathfrak{N}(M)$ can be found for which $\theta'_{\mathfrak{T}}(\theta_0) \neq 0$. Fix $\theta_1 \in \mathfrak{M}$ and let $\theta_n = \theta_0 + \frac{\theta_1}{n}$. Then $(M\theta_n)\theta_n = \frac{1}{n^2}(M\theta_1)\theta_1$ and $\theta'(\theta_n) = \theta'_{\mathfrak{R}(m')}(\theta_0) + \theta'_{\mathfrak{T}}(\theta_0) + \theta'\left(\frac{\theta_1}{n}\right)$. Since $\theta'_{\mathfrak{R}(m')} = m'h$ for some $h \in H(B)$ one has $\theta'_{\mathfrak{R}(m')}(\theta_0) = (h, m(\theta_0))_\xi$. The latter expression is zero by virtue of its being in $\mathfrak{N}(M)$, so

$$\frac{(\theta'(\theta_n))^2}{(M\theta_n)\theta_n} = n^2\left[\frac{(\theta'_{\mathfrak{T}}(\theta_0))^2}{(M\theta_1)\theta_1} + O(n^{-1})\right]$$

and the theorem is proved. □

EXAMPLE 2.1. With reference to the introduction, let the unknown variable force be $\{\theta(t) : t \in [0, 1]\}$. The observation is $\{Y(x, t) : t \in [0, 1]\}$ when mass $x \in [a, b]$ is used, where $Y(t) = \int_0^1 (t-u)_+ \frac{\theta(u)}{x}\, du + N(t)$ and $N(t)$ is a zero mean process with covariance $K(s, t) = \int_0^1 (t-u)_+(s-u)_+\, du$. As usual $a_+ = \max\{0, a\}$. The quantity to be estimated is $\int_0^1 \theta(s)\, ds = \tau(\theta)$. The space $H(K)$ is the set of all functions f on $[0, 1]$ for which $f(0) = f'(0) = 0$ and $f'' \in L_2[0, 1]$. The inner product is

$$(f, g)_K = \int_0^1 f''(s) g''(s)\, ds.$$

For any $\theta \neq 0$ in $\Theta = C[0, 1]$ and $\xi \in \Xi$

$$\begin{aligned}(M(\xi)\theta)\theta &= \Sigma \xi(x_j) \|m_{x_j}(\theta)\|_K^2 \\ &= \Sigma \frac{\xi(x_j)}{x_j^2} \int_0^1 \theta^2(s)\, ds > 0\end{aligned}$$

so

$$d(\tau, \xi) = \sup_{\theta \neq 0} \frac{\left(\int_0^1 \theta(s)\, ds\right)^2}{\Sigma_{j=1}^r \frac{\xi(x_j)}{x_j^2} \int_0^1 \theta^2(s)\, ds}.$$

From the Schwarz inequality one has

$$d(\tau, \xi) = \frac{1}{\Sigma \frac{\xi(x_j)}{x_j^2}}$$

so the optimal design takes all observations at $x = a$.

EXAMPLE 2.2. Assume the same conditions as in example 2.1 except that we wish to estimate $\theta(\frac{1}{2})$. We shall show that $\theta(\frac{1}{2})$ is not estimable with respect to any design $\xi \in \Xi$.

First we observe that for any $\xi \in \Xi$ and $\theta \in C[0, 1]$

$$\begin{aligned}\|m_\xi(\theta)\|_\xi^2 &= \Sigma_{x \in S(\xi)} \xi(x) \int_0^1 \frac{\theta^2(s)}{x^2}\, ds \\ &\leqslant \left(\Sigma \frac{\xi(x)}{x^2} \|\theta\|_\infty^2\right)\end{aligned}$$

so that m_ξ is continuous from $C[0, 1]$ to $H(K)$. Thus m'_ξ is actually the adjoint m_ξ^*

mapping $H(K)$ into the topological dual of $C[0, 1]$. Therefore, given $v \in H(K)$ there is a function g of bounded variation on $[0, 1]$ such that for all $\theta \in C[0, 1]$

$$(m_\xi^* v)\theta = \int_0^1 \theta(s)\, dg(s).$$

If $\theta(\frac{1}{2})$ were estimable one would have $v \in H(K)$ such that $m_\xi^* v$ corresponds to

$$\begin{aligned} g(s) &= 0 \qquad s < \tfrac{1}{2} \\ &= 1 \qquad \tfrac{1}{2} \leqslant s. \end{aligned} \tag{12}$$

However, using the facts that

$$\begin{aligned} (m_\xi^* v)\theta = (v, m_\xi(\theta)) &= \Sigma_{x \in S(\xi)} \xi(x) \int_0^1 v''(s) \frac{\theta(s)}{x}\, ds \\ &= \left(\Sigma \frac{\xi(x)}{x}\right) \int_0^1 v''(s)\theta(s)\, ds \end{aligned} \tag{13}$$

and $\int_0^1 (v''(s))^2\, ds < \infty$ it is clear that (13) implies $m_\xi^* v$ can not correspond to the g in (12).

3. Characterization of optimal designs. In this section we shall use certain assumptions.

(A1) The parameter space Θ is a real Hilbert space.

(A2) The mappings $\{m_x\}_{x \in X}$ from Θ to $H(K)$ are all linear and continuous and $\mathfrak{R}(m_\xi)$ is closed for $\xi \in \Xi$. Since for each $\xi \in \Xi$ the assumption (A2) implies that m_ξ' has values in Θ^*, the topological dual of Θ, we shall write m_ξ^* rather than m_ξ'. Let $\mathfrak{F}$ be the set of all maps ϕ from X into $H(K)$ for which $\|\phi(x)\| \leqslant 1$ for all x and

$$\mathfrak{R} = \{\int m_x^* \phi(x)\, d\xi(x) : \xi \in \Xi, \phi \in \mathfrak{F}\}.$$

(A3) There is a proper closed supporting hyperplane to $\mathfrak{R}$ at each of its boundary points.

(A4) For each $\theta \in \Theta$, $\theta \neq 0$, $\sup_X \|m_x(\theta)\| > 0$.

Results corresponding to those below have been obtained for Θ a linear topological space in Spruill (1978).

Since Hilbert spaces are reflexive, we shall as usual identify Θ^* with Θ. Fix $\theta^* \in \Theta$ and let $v_0 = \inf_\Xi d(\theta^*, \xi)$.

LEMMA 3.1. *Let* A1–A4 *hold. Then*

(a) $\beta\theta^* \in \mathfrak{R}$ *implies* $v_0 \leqslant \dfrac{1}{\beta^2}$, *and*

(b) $\beta\theta^* \in \partial\mathfrak{R}$ *implies* $v_0 \geqslant \dfrac{1}{\beta^2}$.

PROOF. (a) If $\beta = 0$ the result is trivially true. Otherwise

$$\beta\theta^* = \Sigma_{j=1}^n \alpha_j m_{x_j}^* \phi_{x_j}.$$

One has for any $\theta \in \Theta$ that

$$(\theta^*, \theta)^2 = \frac{[\beta(\theta^*, \theta)]^2}{\beta^2} = \frac{[\Sigma \alpha_j(\phi_{x_j}, m_{x_j}(\theta))_K]^2}{\beta^2}$$
$$\leqslant \frac{\Sigma \alpha_j(M_{x_j}\theta, \theta)}{\beta^2}$$
$$= \frac{(M(\xi)\theta, \theta)}{\beta^2}.$$

Therefore, by lemma 2.7, $v_0 \leqslant \frac{1}{\beta^2}$.

(b) Since $\beta\theta^* \in \partial\mathfrak{R}$ there is a $\theta \in \Theta$, $\theta \neq 0$, such that $\beta(\theta^*, \theta) \geqslant |(r, \theta)|$ for all $r \in \mathfrak{R}$. This is a consequence of (A3). Let $x_n \in X$ be such that $\|m_{x_n}(\theta)\| \to \sup_X \|m_x(\theta)\|$. Let

$$r_n = m^*_{x_n}\left(\frac{m_{x_n}(\theta)}{\|m_{x_n}(\theta)\|}\right)$$

to obtain

$$\beta(\theta^*, \theta) \geqslant |(r_n, \theta)| = \|m_{x_n}(\theta)\|. \tag{14}$$

We conclude that $\beta(\theta^*, \theta) \geqslant \sup_X \|m_x(\theta)\|$. Also, since $\mathfrak{R}$ is symmetric and convex, for $\varepsilon > 0$ sufficiently small $(\beta - \varepsilon)\theta^* \in \mathfrak{R}$ ($\beta > 0$ by (14) and (A4)). One has $(\beta - \varepsilon)\theta^* = \Sigma \alpha_j m^*_{x_j}(\phi_j)$ so $\beta(\theta^*, \theta) - \varepsilon(\theta^*, \theta) = \Sigma \alpha_j(\phi_j, m_{x_j}(\theta))$

$$\leqslant \Sigma \alpha_j \|m_{x_j}(\theta)\| \leqslant \sup_X \|m_x(\theta)\|. \tag{15}$$

Combining (14) and (15) we have

$$\beta(\theta^*, \theta) = \sup_X \|m_X(\theta)\|. \tag{16}$$

Let ξ_n be such that $d(\theta^*, \xi_n) \to v_0$. If v_0 is $+\infty$ the result is trivially true. Assume, therefore that $\theta^* \in \mathfrak{R}(m^*_{\xi_n})$ for each n. Then θ must be such that $(M(\xi_n)\theta, \theta) > 0$ for all n. To see this, suppose that it is not for some n. Then $\theta \in \mathfrak{N}[M(\xi_n)] = \mathfrak{R}^\perp[M'(\xi_n)]$, and since $M(\xi_n) = M'(\xi_n)$ (see lemma 2.7) one has $\theta \in \mathfrak{R}^\perp[M(\xi_n)]$. From the argument in lemma 2.7 there is a sequence $\theta^*_k \in \mathfrak{R}[M(\xi_n)]$ such that $\lim_{k\to\infty}|d(\theta^*_k, \xi_n) - d(\theta^*, \xi_n)| = 0$. But $\theta^*_k(\theta) \equiv 0$ by definition of $\mathfrak{R}^\perp[M(\xi_n)]$ so that $\theta^*(\theta) = 0$. This, with (16) contradicts (A4). Since $(M(\xi_n)\theta, \theta) > 0$ for all n one has

$$\frac{1}{\beta^2} = \frac{(\theta^*, \theta)^2}{\sup_X \|m_x(\theta)\|^2_K} \leqslant \frac{(\theta^*, \theta)^2}{(M(\xi_n)\theta, \theta)} \leqslant d(\theta^*, \xi_n). \qquad \square$$

THEOREM 3.1. *Let* (θ^*, θ) *be estimable with respect to some design and* A1–A4 *hold. The design* ξ_0 *is optimal for the estimation of* (θ^*, θ) *if and only if there is a function* $\phi : X \to H(K)$ *such that* $\|\phi(x)\| \equiv 1$ *and*

$$\int m^*_x \phi(x)\, d\xi_0(x)$$

is (i) *proportional to* θ^* *and* (ii) *in* $\mathfrak{R} \cap \partial\mathfrak{R}$.

PROOF. Suppose ξ_0 is optimal. Then setting $\lambda_0 = M^{\#}(\xi_0)\theta^*$ one has $M_0\lambda_0 = \theta^*$ and $(M_0\lambda_0, \lambda_0) = v_0$. Let $\{z : (\lambda, z - v_0^{-\frac{1}{2}}\theta^*) = 0$ be a supporting hyperplane to $\mathfrak{R}$ at $v_0^{-\frac{1}{2}}\theta^*$ where $\lambda \neq 0$. From above $(\theta^*, v_0^{-\frac{1}{2}}\lambda) = \sup_X \|m_x(\lambda)\|$. Since $M_0\lambda_0 = \theta^*$ one has $(v_0^{-\frac{1}{2}}M_0\lambda_0, \lambda) = \sup_X \|m_x(\lambda)\|$. Therefore

$$\begin{aligned} \sup\|m_x(\lambda)\|^2 &= v_0^{-1}(M_0\lambda_0, \lambda)^2 \\ &\leqslant v_0^{-1}\big(M_0(M_0^{\#}\theta^*), M_0^{\#}\theta^*\big)(M_0\lambda, \lambda) \qquad (17)\\ &= v_0^{-1}\big(\theta^*, M_0^{\#}\theta^*\big)(M_0\lambda, \lambda) = (M_0\lambda, \lambda) \end{aligned}$$

with strict inequality unless $m_{\xi_0}(\lambda_0) = km_{\xi_0}(\lambda)$ for some scalar k. Since one always has $(M_0\lambda, \lambda) \leqslant \sup_X\|m_x(\lambda)\|^2$ equality holds in (17) and we have $\|m_x(\lambda)\| = \sup_X\|m_x(\lambda)\|$ on the support $S(\xi_0)$ of ξ_0. Set $\phi(x) = \dfrac{m_x(\lambda)}{\|m_x(\lambda)\|}$ for $x \in S(\xi_0)$. Then

$$\int m_x^*\phi(x)\, d\xi_0(x) = \frac{M_0\lambda}{\sup\|m_x(\lambda)\|} = \frac{M_0\lambda_0}{k \sup\|m_x(\lambda)\|}.$$

From above

$$\begin{aligned} v_0 &= (M_0\lambda_0, \lambda_0) = k(M_0\lambda, \lambda_0) = k((M_0\lambda_0, \lambda_0))^{\frac{1}{2}}((M_0\lambda, \lambda))^{\frac{1}{2}} \\ &= kv_0^{\frac{1}{2}} \sup_X\|m_x(\lambda)\|, \end{aligned}$$

so that $\int m_x^*\phi(x)\, d\xi_0(x) = v_0^{-\frac{1}{2}}\theta^*$.

If (i) and (ii) hold then for any $\theta \in \Theta$

$$\begin{aligned} v_0^{-1}(\theta^*, \theta)^2 &= \big(\int (m_x(\theta), \phi(x))_k\, d\xi_0(x)\big)^2 \\ &\leqslant \int \|m_x(\theta)\|^2\, d\xi_0(x) = (M_0\theta, \theta) \end{aligned}$$

showing ξ_0 to be optimal. □

There are other theorems of possible interest. One of these, whose proof we leave to the reader (see Studden and Tsay (1976)), is the following. Let $\theta_0^* \in \Theta$ be fixed and $\Delta = \{\delta \in \Theta : (\theta_0^*, \delta) = 1\}$.

THEOREM 3.2. *Under* A1–A4, *if* $\delta_0 \in \Delta$ *satisfies*

(i) $\inf_\Delta \sup_X\|m_x(\delta)\| = \sup_X\|m_x(\delta_0)\|$,

(ii) $S(\xi_0) \subset \{x : \|m_x(\delta_0)\| = \sup_X\|m_x(\delta_0)\|\}$, *and*

(iii) $\int m_x^* m_x(\delta_0)\, d\xi_0(x)$ *is proportional to* θ_0^* *then* ξ_0 *is optimal for estimating* θ_0^* *and* $v_0 = (\sup_X\|m_x(\delta_0)\|)^{-2}$. *Furthermore, if there is an optimal design* ξ_0, *then there is a* δ_0 *satisfying* (i), (ii), *and* (iii).

4. Some examples. Verification of assumption (A3) may be a nontrivial task unless $\Theta = R^n$ for some n. We digress briefly in order to provide an easily verified sufficient condition for (A3).

The point a is said to be an *internal* (not necessarily interior) point of the set $A \subset \Theta$ if $a \in A$ and for each $\theta \in \Theta$ there is an $\alpha_0 > 0$ such that $\alpha\theta + (1 - \alpha)a \in A$ for $0 \leqslant \alpha < \alpha_0$. Since $\overline{\mathfrak{R}}$ is symmetric, $\overline{\mathfrak{R}}$ has an internal point if and only if 0 is an internal point. Furthermore, since Θ is a Banach space, an application of the Baire category theorem shows that $\overline{\mathfrak{R}}$ has an internal point if and only if it has an interior point. It is well known (see, for example, Holmes (1975)) that if $A \subset \Theta$ is convex and $\overline{A}$ contains an interior point then there is a supporting hyperplane to A at each of its boundary points.

THEOREM 4.1. *If m_x is continuous for each $x \in X$, there is a constant $k > 0$ such that for all $\theta \in \Theta$*

$$\sup_X \|m_x(\theta)\| \geqslant k\|\theta\|, \tag{18}$$

and (A1) *holds, then* (A3) *holds.*

PROOF. It suffices to prove that 0 is an internal point of $\overline{\mathfrak{R}}$. Let $\theta_0 \in \Theta$ be nonzero and suppose $\alpha\theta_0$ is not in $\overline{\mathfrak{R}}$. Then there is a $\lambda_\alpha \in \Theta$ such that $\|\lambda_\alpha\| = 1$, and $(\alpha\theta_0, \lambda_\alpha) \geqslant a(\alpha) > b(\alpha) \geqslant (r, \lambda_\alpha)$ for all $r \in \overline{\mathfrak{R}}$ (see Dunford and Schwarz (1959) V.2.10). Therefore

$$\alpha\|\theta_0\| \geqslant \alpha(\theta_0, \lambda_\alpha) > \|m_x(\lambda_\alpha)\| \tag{19}$$

for all $x \in X$ such that $\|m_x(\lambda_\alpha)\| > 0$. This may be seen by choosing $r = m_x^* \dfrac{m_x(\lambda_\alpha)}{\|m_x(\lambda_\alpha)\|}$. In view of our assumptions, (19) implies 0 is an internal point of $\overline{\mathfrak{R}}$.

EXAMPLE 4.1. Let $X = [0, 2\pi]$, $\Theta = L_2[0, 1]$, and

$$Y(x, t) = \int_0^t \theta(s)\sin(2\pi s - x)\, ds + W(t)$$

for $t \in [0, 1]$. Since

$$\|m_x(\theta)\|^2 = \int_0^1 \theta^2(t) \sin^2(2\pi t - x)\, dt \leqslant \|\theta\|_2^2$$

(A1) and (A2) are satisfied. Let $\phi(x) = \frac{1}{2\pi}\chi(x)_{[0, 2\pi]}$. Then

$$\sup_{0 \leqslant x \leqslant 2\pi}\|m_x(\theta)\|^2 \geqslant \int_0^{2\pi}\|m_x(\theta)\|^2\phi(x)\, dx = \frac{\|\theta\|_2^2}{2}$$

so (A3) and (A4) are satisfied. We shall use Theorem 3.2 to solve the design problem for estimating $\int_0^1\theta(s)\, ds$. If we view X as the parameter space [0, 1] as the outcome space, $\mathfrak{R}$ as the action space, and $\theta(\cdot)$ a decision rule in a statistical decision problem then $\theta_0(s) \equiv 1$ results in a constant risk of $\frac{1}{2}$. Viewing ϕ as a prior on the parameter space, it is easily seen that θ_0 is Bayes among those $\theta \in L_2$ such that $\int_0^1\theta(s)\, ds = 1$. That is; it is minimax,

$$\inf_{\theta:\{\int_0^1\theta(s)\, ds = 1\}}\sup_{0 \leqslant x \leqslant 2\pi}\|m_x(\theta)\|^2 = \sup_{0 \leqslant x \leqslant 2\pi}\|m_x(\theta_0)\|^2 = \tfrac{1}{2}.$$

It remains to find a design for which

$$\left[\Sigma\xi(x)m_x^* m_x(\theta_0)\right](t) \equiv 1. \tag{20}$$

Since $m_x^* m_x(\theta_0)(t) = \theta_0(t)\sin^2(2\pi t - x)$ the design which places masses $\frac{1}{2}$ at $x = 0$ and $\frac{1}{2}$ at $x = \pi/2$ satisfies (20) and is optimal for estimating $\int_0^1\theta(s)\, ds$ with $v = 2$.

We close with an example in which $\mathfrak{R}$ is the convex hull of a closed bounded set but an estimable functional fails to have an associated optimal design.

EXAMPLE 4.2. Let $\Theta = l_2$, T be a one point set $x \in l_2$ and

$$Y(x) = (\theta, x) + \varepsilon.$$

Let $X = \{\pm(\Phi_0 + \Phi_j) : j \geqslant 1\} \cup \left\{\frac{\pm\Phi_0}{2}\right\}$ where $\{\Phi_0, \Phi_1, \cdots\}$ is a complete orthonormal system in l_2. Since $m_x^* u = ux$ one finds that $\mathfrak{R} = co(X)$. We observe that (Φ_0, θ) is estimable since $m_{\Phi_0/2}^*(2) = \Phi_0$. Also $\Phi_0 \in \overline{\mathfrak{R}}$ since taking $\alpha_{n_j} = \frac{1}{n}$, $j = 1, \cdots, n$ one has

$$\|\Phi_0 + \Sigma_{j=1}^n \alpha_{n_j} \Phi_j - \Phi_0\|^2 = \frac{1}{n} \to 0.$$

It is easily seen that Φ_0 is not in $\mathfrak{R}$ so that there is no optimal design for estimating (Φ_0, θ).

5. Acknowledgement. I would like to thank S. Demko and H. Engl for their helpful conversations.

REFERENCES

DUNFORD, N. and SCHWARTZ, J. T. (1958). *Linear Operators Part I. General Theory.* Interscience, New York.

FEDOROV, V. V. (1972). *Theory of Optimal Experiments.* Academic, New York.

HOLMES, R. B. (1975). *Geometric Functional Analysis and its Applications.* Springer, New York.

KIEFER, J. (1974). General equivalence theory for optimum designs (approximate theory). *Ann. Statist.* **2** 849–879.

MEHRA, RAMAN K. (1974). Optimal input signals for parameter estimation in dynamic systems survey and new results IEEE Trans. Automat. Contr. AC-19, 753–768.

NASHED, M. ZUHAIR and VOTRUBA, G. F. (1976). A unified operator theory of generalized inverses. In *Generalized Inverses and Applications.* (M. Z. Nashed, ed.) Academic, New York.

PARZEN, E. (1959). Statistical inference on time series by Hilbert space methods, I. Technical Report No. 23, Stanford University.

PAZMAN, ANDREJ (1978). Hilbert-space methods in experimental design. *Kybernetika* **14** 73–84.

SPRUILL, M. C. (1978). Optimal designs for second order processes with general linear means. Technical Report, School of Mathematics, Georgia Inst. Technology.

SPRUILL, CARL and STUDDEN, W. J. (1978). Optimum designs when the observations are second order processes. *J. Multivariate Anal.* **8** 153–172.

STUDDEN, WILLIAM J. and TSAY, JIA-YEONG (1976). Remez's procedure for finding optimal designs. *Ann. Statist.* **4** 1271–1279.

TAYLOR, A. E. (1958). *Introduction to Functional Analysis.* Wiley, New York.

WAHBA, GRACE (1976). On the optimal choice of nodes in the collocation-projection method for solving linear operator equations. *J. Approximation Theory.* **16** 175–186.

DEPARTMENT OF MATHEMATICS
GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA, GEORGIA 30332