The Annals of Statistics
1979, Vol. 7, No. 5, 1078–1085

DESIGN OF OPTIMAL CONTROL FOR A REGRESSION PROBLEM

BY DER-SHIN CHANG

National Tsing Hua University

Consider the realization of the process $y(t) = \Sigma_{k=1}^{n}\theta_k f_k(t) + \xi(t)$ on the interval $T = [0, 1]$ for functions $f_1(t), f_2(t), \cdots, f_n(t)$ in $H(R)$, the reproducing kernel Hilbert space with reproducing kernel $R(s, t)$ on $T \times T$, where $R(s, t) = E\xi(s)\xi(t)$ is assumed to be continuous and known. Problems of the selection of functions $\{f_k(t)\}_{k=1}^{n}$ are discussed for D-optimal, A-optimal and other criteria of optimal designs.

Introduction. Consider a regression model

$$(1) \qquad y(t) = \Sigma_{k=1}^{n}\theta_k f_k(t) + \xi(t), \qquad t \in T, T = [0, 1]$$

with the noise process $\xi(t)$ having zero mean and known continuous covariance kernel $R(s, t) = E\xi(s)\xi(t)$, $(s, t) \in T \times T$. Let $H(R)$ be the reproducing kernel Hilbert space (RKHS) with reproducing kernel (RK) $R(s, t)$ on $T \times T$, and $\{f_1(t), \cdots, f_n(t)\}$ be a linearly independent set of functions in $H(R)$. Then, by the Gauss-Markov theory of continuous time series (see [8]), we obtain for $f' = (f_1(t), f_2(t), \cdots, f_n(t))$ the minimum variance unbiased estimate $\hat{\theta} = M^{-1}(f)$ $(\langle y, f_1\rangle \sim, \cdots, \langle y, f_n\rangle \sim)'$, and its convariance matrix $\mathrm{Cov}[\hat{\theta}] = M^{-1}(f)$, where $\hat{\theta}' = (\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_n)$, $M(f) = [m_{ij}]_{i,j=1}^{n}$, $m_{ij} = \langle f_i, f_j\rangle_R$ if $\{f_k(t)\}_{k=1}^{n} \subset H(R)$, and $\langle y, f_k\rangle \sim$, $k = 1, 2, \cdots, n$, are defined as if $y(t)$ is an element in $H(R)$. In [1], the author has calculated the optimal functions $\{f_k(t)\}_{k=1}^{n}$ in some special set X of functions to minimize $\mathrm{Var}(\Sigma_{i=1}^{n}a_i\hat{\theta}_i)$ and $\Sigma_{i=1}^{n}\mathrm{Var}(\hat{\theta}_i)$. This problem is similar to that of optimal design of input signals for parameter estimation in automatic control (see [5] and [6]). However, full investigation of model (1) in RKHS and the analytic form of optimal solutions of $\{f_k(t)\}_{k=1}^{n}$ were not given.

In this paper, we give some natural criteria for optimal designs which generalize the idea of Kiefer and Wolfowitz (see [4] and [3]) to continuous realization of $y(t)$ if it is possible to select the functions $\{f_k(t)\}_{k=1}^{n}$ from a set of functions $X \subset H(R)$ prior to the experiment; and discuss some problems of D-optimal, A-optimal and weighted optimal designs and their respective solutions. In Section 1, we give criteria for designs of regression model (1). In Section 2, we solve the D-optimal, A-optimal and weighted optimal design problems and give the optimal solutions of $\{f_k(t)\}_{k=1}^{n}$ in each case. In Section 3, some examples and special cases of Section 2 are discussed.

1. Design criteria. If (1) is given, then it is well known (see [8]) that the space of functions generated by $\{R_t(\cdot), t \in T | R_t(t') = R(t', t)\}$ is a RKHS, denoted by

Received January 1978; revised April 1978.

AMS 1970 *subject classifications.* Primary 62K05; secondary 93E20.

Key words and phrases. Regression model, continuous sense of Gauss-Markov theory, reproducing kernel Hilbert space, continuous sense of D-optimal, A-optimal, weighted optimum design.

$H(R)$, with $R(s, t)$ on $T \times T$, where $R(s, t)$ is symmetric and positive definite, and by Mercer's theorem (see [9], pages 242–246), we know there exists a set of orthonormal functions $\{\phi_v(t)\}_{v=1}^{\infty}$ in $\mathcal{L}^2[T]$ and a corresponding sequence of positive real numbers $\{\lambda_v\}_{v=1}^{\infty}$ such that

(2)
$$R(s, t) = \Sigma_{v=1}^{\infty}\lambda_v\phi_v(s)\phi_v(t)$$

is uniformly convergent in $T \times T$ if $R(s, t)$ is continuous; also that the inner product in $H(R)$ is

$$\langle g, h\rangle_R = \Sigma_{v=1}^{\infty}\frac{g_v h_v}{\lambda_v},$$

where $g_v = (g, \phi_v)_{\mathcal{L}2}$, $h_v = (h, \phi_v)_{\mathcal{L}2}$, for any $g, h \in H(R)$. That is, $H(R) = \{h|\Sigma_{v=1}^{\infty}h_v^2/\lambda_v < \infty, h_v = (h, \phi_v)_{\mathcal{L}2}\}$.

Assume further that a set of linearly independent functions $\{f_k(t)\}_{k=1}^{n}$ in $H(R)$ is given. Then, by [8], we have for $\theta' = (\theta_1, \cdots, \theta_n)$ and $f' = (f_1(t), \cdots, f_n(t))$ the minimum variance unbiased estimate

(3)
$$\hat{\theta} = M^{-1}(f)(\langle y, f_1\rangle \sim, \cdots, \langle y, f_n\rangle \sim)'$$

with $\text{Cov}[\hat{\theta}] = M^{-1}(f)$ and

$$\langle y, f_k\rangle \sim = \Sigma_{v=1}^{\infty}\frac{f_{kv}}{\lambda_v}y_v, \qquad k = 1, 2, \cdots, n,$$

with $y_v = (y, \phi_v)_{\mathcal{L}2}$, the stochastic integral of $y(t)$ with respect to weight function $\phi_v(t) \in \mathcal{L}^2[T]$ (see [7] and [8]).

Now extending the idea of [4] and [2], we can define D-optimal, A-optimal and weighted optimal design as follows.

DEFINITION. In model (1), for an experiment with $f^* = (f_1^*(t), \cdots, f_n^*(t))'$ and $X \subset H(R)$ if

(i) $\max_{\{f_k\}_{k=1}^n \subset X}|M(f)| = |M(f^*)|$, it is called a D-optimal design in set X;

(ii) $\min_{\{f_k\}_{k=1}^n \subset X}\text{tr } M^{-1}(f) = \text{tr } M^{-1}(f^*)$, it is called an A-optimal design in set X;

(iii) $\min_{\{f_k\}_{k=1}^n \subset X}\text{tr } WM^{-1}(f) = \text{tr } WM^{-1}(f^*)$, it is called a weighted optimal design in set X for weighting matrix W.

The reasons that we take these criteria for designs are direct extensions of those in standard optimal design work (see [2] and [4]). For instance, if $W = aa'$, $a = (a_1, a_2, \cdots, a_n)'$, then $\text{tr } WM^{-1}(f) = \text{Var}[\Sigma_{i=1}^n a_i\hat{\theta}_i]$, so that the minimization of (iii) has the statistical meaning which is discussed in other space of functions in [1]. Further, the reasons that we restrict X, our design space of functions, are based on practical considerations (see [5] and [6]).

2. Main results.

In this section we state and prove the following

THEOREM 1. *Suppose model* (1) *is given, and* $X = \{h|h \in H(R), \|h\|_R^2 \leq L\}$; L *is any given positive number. Then under the assumptions in Section* 1,

(i) $\max_{\{f_n\}_{k=1}^n \subset X} |M(f)| = L^n$, *which is attainable at* $f_k^*(t) = (L\lambda_k)^{\frac{1}{2}}\phi_k(t)$, $k = 1, 2, \cdots, n$;

(ii) $\min_{\{f_k\}_{k=1}^n \subset X} \operatorname{tr} M^{-1}(f) = n/L$, *which is attainable at* $f_k^*(t) = (L\lambda_k)^{\frac{1}{2}}\phi_k(t)$, $k = 1, 2, \cdots, n$;

(iii) $\min_{\{f_k\}_{k=1}^n \subset X} \operatorname{tr} WM^{-1}(f) = \Sigma_1^n \eta_i / L$, *if W is a known symmetric and positive definite matrix with eigenvalues* $\eta_1 \geqslant \eta_2 \geqslant \cdots \geqslant \eta_n > 0$.

PROOF. (i) Since $M(f)$ is positive definite, it is well known (Hadamard's inequality) that $|M(f)| \leqslant \Pi_{k=1}^n m_{kk} = \Pi_{k=1}^n \Sigma_{j=1}^\infty f_{kj}^2/\lambda_j$. Then, since $\{f_k\}_{k=1}^n \subset X$, that is, $\Sigma_{j=1}^\infty f_{kj}^2/\lambda_j \leqslant L$ for $k = 1, 2, \cdots, n$, we have

$$|M(f)| \leqslant \Pi_{k=1}^n L = L^n.$$

The maximum is attainable at the boundary of X. For example, we can take

$$f_k^*(t) = (L\lambda_k)^{\frac{1}{2}}\phi_k(t), \qquad k = 1, 2, \cdots, n. \tag{4}$$

Note that there are many D-optimal designs, the set of (4) being just an easy one to select.

(ii) Suppose first that $\{f_k\}_{k=1}^n$ be a set of linearly independent functions with finite square norms, say $\|f_k\|_R^2 = L_k$, $k = 1, 2, \cdots, n$. Then, by the Gram-Schmidt process, we can get an orthonormal collection $u_1(t), u_2(t), \cdots, u_n(t)$ in $H(R)$ such that

$$\begin{aligned} u_i(t) &= a_{i1} f_1(t) + \cdots + a_{ii} f_i(t) \quad \text{and} \\ f_i(t) &= b_{i1} u_1(t) + \cdots + b_{ii} u_i(t), \qquad i = 1, 2, \cdots, n. \end{aligned} \tag{5}$$

Let

$$A = \begin{bmatrix} a_{11} & & & 0 \\ a_{21} & a_{22} & & \\ \vdots & & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} b_{11} & & & 0 \\ b_{21} & b_{22} & & \\ \vdots & & \ddots & \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix}$$

Then, for $u' = (u_1(t), \cdots, u_n(t))$ and $f' = (f_1(t), \cdots, f_n(t))$, (5) can be rewritten as

$$u = Af \quad \text{and} \quad f = Bu. \tag{6}$$

Further, by (6), we have $B = A^{-1}$. Now since

$$M(f) = [m_{ij}]_{i,j=1}^n, \qquad m_{ij} = \langle f_i, f_j \rangle_R,$$

then, by (5), we have

$$m_{ij} = \langle \Sigma_{k=1}^i b_{ik} u_k, \Sigma_{l=1}^j b_{jl} u_l \rangle_R = \Sigma_{l=1}^{\min(i,j)} b_{il} b_{jl} = (BB')_{ij}.$$

Thus,

$$M(f) = BB' = (A'A)^{-1},$$

so

(7) $$M^{-1}(f) = A'A.$$

Hence

(8) $$\operatorname{tr} M^{-1}(f) = \operatorname{tr} A'A = \Sigma_{i=1}^{n}\Sigma_{j=i}^{n} a_{ij}^2 \geqslant \Sigma_{i=1}^{n} a_{ii}^2.$$

Equality in (8) occurs only if $a_{ij} = 0$ for all $i \neq j$; that is,

$$\langle f_i, f_j \rangle_R = 0 \qquad \forall i \neq j.$$

Therefore,

$$\min_{\|f_k\|_R^2 = L_k,\, k=1,2,\cdots,n} \operatorname{tr} M^{-1}(f) = \Sigma_{k=1}^{n} a_{kk}^2 = \Sigma_{k=1}^{n} \frac{1}{\|f_k\|_R^2} = \Sigma_{k=1}^{n} \frac{1}{L_k}.$$

Finally, since $\{f_k(t)\}_{k=1}^{n}$ must be in X,

$$\min_{\{f_k\}_{k=1}^{n} \subset X} \operatorname{tr} M^{-1}(f) = \min_{L_k \leqslant L,\, k=1,2,\cdots,n} \Sigma_{k=1}^{n} \frac{1}{L_k} = n/L,$$

which is attainable for any collection of orthogonal functions $\{f_k(t)\}_{k=1}^{n}$ with square norms equal to L. For example, $f_k^*(t) = (L\lambda_k)^{\frac{1}{2}}\phi_k(t)$, $k = 1, 2, \cdots, n$, is an A-optimal design.

(iii) If W is a symmetric and positive definite matrix, then there exists an orthogonal matrix P and diagonal matrix D such that $W = PDP'$, with

$$D = \begin{bmatrix} \eta_1 & & 0 \\ & \ddots & \\ 0 & & \eta_n \end{bmatrix}, \qquad \eta_1 \geqslant \eta_2 \geqslant \cdots \geqslant \eta_n > 0.$$

By the proof in (ii) and (7),

$$M^{-1}(f) = A'A.$$

Thus,

$$\operatorname{tr} WM^{-1}(f) = \operatorname{tr} PDP'A'A = \operatorname{tr} DP'A'AP = \operatorname{tr} D\tilde{M}^{-1}(g)$$
$$= \Sigma_{i=1}^{n} \eta_i (\tilde{M}^{-1})_{ii},$$

where $g = (g_1(t), \cdots, g_n(t))' = P'f$ and $\tilde{M}^{-1}(g) = \tilde{A}'\tilde{A}$ with

$$\tilde{A} = AP = \begin{bmatrix} A_{11} & & & \\ & & 0 & \\ A_{21} & A_{22} & & \\ \vdots & & \ddots & \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}, \text{ by (6)},$$

and

$$A_{11} = 1/\| g_1\|_R,$$

$$A_{21} = -\langle g_2, g_1\rangle_R \frac{1}{\| g_1\|_R^2 \; \| g_2 - \langle g_2, g_1\rangle_R g_1(1/\| g_1\|_R^2)\|_R},$$

$$A_{22} = 1/\| g_2 - \langle g_2, g_1\rangle_R g_1(1/\| g_1\|_R^2)\|_R, \text{ and so on.}$$

Hence

$$(\tilde{M}^{-1})_{ii} = \Sigma_{k=i}^n A_{ki}^2,$$

and

$$\operatorname{tr} WM^{-1}(f) = \Sigma_{i=1}^n \eta_i \Sigma_{k=i}^n A_{ki}^2 \geqslant \Sigma_{i=1}^n \eta_i A_{ii}^2 \geqslant \Sigma_{i=1}^n \eta_i \frac{1}{\| g_i\|_R^2}$$

$$\geqslant \Sigma_{i=1}^n \eta_i / L.$$

The lower bound is attainable if $\langle g_i, g_j\rangle_R = 0$ for all $i \neq j$, and $\| g_i\|_R^2 = L, i = 1, 2, \cdots, n$. Thus, since $f = Pg$, P orthogonal,

$$\begin{aligned} \langle f_i, f_j\rangle_R &= 0 \qquad i \neq j \\ &= L \qquad i = j. \end{aligned}$$

Therefore,

$$\min_{\{f_k\}_{k=1}^n \subset X} \operatorname{tr} WM^{-1}(f) = \Sigma_{i=1}^n \eta_i / L,$$

and the minimum is attainable at any set of orthogonal functions with square norm equal to L. □

COROLLARY. *The design* $\{f_k^*(t)\}_{k=1}^n, f_k^*(t) = (L\lambda_k)^{\frac{1}{2}}\phi_k(t)$, *is simultaneously D- and A-optimal, and furthermore is weighted optimal for any symmetric positive definite matrix W.*

This result follows trivially from Theorem 1, and is similar to, but a stronger result than, the discrete case (see [2], page 139).

3. Some other optimal designs and examples. In Section 2 we restricted our discussion on weighted optimal design to that for positive definite matrix W. There are many cases of W nonnegative definite in which it is very complicated to construct the optimal solutions $f_k^*(t)$, $k = 1, 2, \cdots, n$. But if $W = aa'$ with $a' = (a_1, a_2, \cdots, a_n)$ and $|a_1| \leqslant |a_2| \leqslant \cdots \leqslant |a_n|$, then we can follow [1] and obtain the following

THEOREM 2. *Suppose* $a_1, a_2, \cdots, a_n$ *are given real numbers such that* $|a_1| \leqslant |a_2| \leqslant \cdots \leqslant |a_n|$. *Let* $X = \{h | h \in H(R), \|h\|_R^2 \leqslant L\}$, $W = aa'$, *where* $a' = (a_1, \cdots, a_n)$. *Then, under model* (1), *we have*

$$\min_{\{f_k\}_{k=1}^n \subset X} \operatorname{tr} WM^{-1}(f) = a_n^2/L,$$

which is attainable if we take $f_n^*(t) = \text{sign } a_n(L^{\frac{1}{2}}/n^{\frac{1}{2}})\Sigma_{i=1}^n(\lambda_i)^{\frac{1}{2}}\,\phi_i(t)$ *and* $f_1^*(t), \cdots, f_{n-1}^*(t)$ *any functions such that* $\langle f_i^*, u\rangle_R = a_i L^{\frac{1}{2}}/|a_n|$, $i = 1, 2, \cdots, n-1$, *where* $u(t) = (1/n^{\frac{1}{2}})\Sigma_{i=1}^n(\lambda_i)^{\frac{1}{2}}\phi_i(t)$.

Proof. Let $\eta = \Sigma_{i=1}^n a_i\theta_i$ and $|a_1| \leqslant |a_2| \leqslant \cdots \leqslant |a_n|$. Then, by (3)

$$\hat{\eta} = \Sigma_{i=1}^n a_i\hat{\theta}_i = a'\hat{\theta} = \Sigma_{k=1}^n c_k\langle y, f_k\rangle \sim,$$

where

$$a' = (a_1, \cdots, a_n), \qquad \theta' = (\theta_1, \theta_2, \cdots, \theta_n)$$

and

$$c_k,\ k = 1, 2, \cdots, n \text{ are functions of } \{m_{ij}\}_{i,j=1}^n.$$

Since $\hat{\theta}$ is the minimum variance unbiased estimate of θ, we have

$$E\hat{\eta} = \Sigma_{i=1}^n a_i E\hat{\theta}_i = \Sigma_{i=1}^n a_i\theta_i = \eta,$$

which in turn implies

$$\Sigma_{k=1}^n c_k E\langle y, f_k\rangle \sim = \eta. \tag{9}$$

Next, by [7] and [8], we have

$$E\langle y, f_k\rangle \sim = \langle \Sigma_{l=1}^n \theta_l f_l, f_k\rangle_R \tag{10}$$

and

$$\text{Var}\langle y, f_k\rangle \sim = \langle f_k, f_k\rangle_R.$$

Thus, (9) can be rewritten as

$$\Sigma_{l=1}^n\theta_l\Sigma_{k=1}^n c_k\langle f_l, f_k\rangle_R = \Sigma_{l=1}^n a_l\theta_l.$$

That is,

$$\Sigma_{k=1}^n c_k\langle f_l, f_k\rangle_R = a_l, \qquad l = 1, 2, \cdots, n. \tag{11}$$

Now, let $g(t) = \Sigma_{k=1}^n c_k f_k(t) = \Sigma_{k=1}^n c_k\Sigma_{j=1}^\infty f_{kj}\phi_j(t) = \Sigma_{j=1}^\infty g_j\phi_j(t)$, where $g_j = (g, \phi_j)_{\mathcal{L}2}, j = 1, 2, \cdots$, so that $\hat{\eta} = \Sigma_{k=1}^n c_k\langle y, f_k\rangle \sim = \langle y, g\rangle \sim$. Thus,

$$\text{Var}[\hat{\eta}] = \text{Var}[\Sigma_{i=1}^n a_i\hat{\theta}_i] = \text{Var}[\langle y, g\rangle \sim]$$
$$= \langle g, g\rangle_R = \sigma^2, \text{ say, by (10).}$$

Finally, we investigate the lower bound of $\sigma > 0$. Since, by (11), we have

$$\langle g, f_l\rangle_R = a_l, \qquad l = 1, 2, \cdots, n.$$

Let $u(t) = (1/\sigma)g(t)$. Then $\langle u, u\rangle_R = 1$, and

$$\sigma\langle u, f_l\rangle_R = a_l, \qquad l = 1, 2, \cdots, n.$$

That is, by linear independence of $\{f_k(t)\}_{k=1}^n$, we have, for $l = 1, 2, \cdots, n$,

$$\sigma = a_l/\langle u, f_l\rangle_R \geqslant a_l/\|f_l\|_R,$$

by Cauchy-Schwartz inequality. Therefore, if $\{f_k(t)\}_{k=1}^n \subset X$, we have

$$\operatorname{tr} WM^{-1}(f) = \operatorname{Var}[\hat{\eta}] = \sigma^2 \geqslant \frac{a_n^2}{L} \geqslant \frac{a_{n-1}^2}{L} \geqslant \cdots \geqslant \frac{a_1^2}{L},$$

which implies that σ^2 cannot be smaller than a_n^2/L, and the optimal choices of $\{f_k(t)\}_{k=1}^n$ are

$$f_n^*(t) = \operatorname{sign} a_n L^{\frac{1}{2}} u(t)$$

with $\langle u, u\rangle_R = 1$ (for example, $u(t) = (1/n^{\frac{1}{2}})\Sigma_{i=1}^n(\lambda_i)^{\frac{1}{2}}\phi_i(t))$, and $\{f_1^*(t), \cdots, f_{n-1}^*(t)\}$are any linearly independent functions satisfying

$$\langle u, f_j^*\rangle_R = a_j L^{\frac{1}{2}}/|a_n|, \qquad j = 1, 2, \cdots, n-1. \quad \square$$

We now give two examples to conclude our discussion.

EXAMPLE 1. Suppose that in the interval [0, 1] it is possible to observe a realization of the process

$$y(t) = \theta_1 f_1(t) + \theta_2 f_2(t) + \xi(t)$$

with $E\xi(t) = 0$ and known continuous covariance function $E\xi(s)\xi(t) = R(s, t) = \Sigma_{i=1}^\infty \lambda_i \phi_i(s)\phi_i(t)$. Then an optimal solution of $\{f_1(t), f_2(t)\}$ to minimize $\operatorname{Var}(a_1\hat{\theta}_1 + a_2\hat{\theta}_2)$ with $|a_1| \leqslant |a_2|$, in the set $X = \{h : \|h\|_R^2 \leqslant L\}$ is:

$$f_2^*(t) = \frac{L^{\frac{1}{2}}}{2^{\frac{1}{2}}}\left((\lambda_1)^{\frac{1}{2}}\phi_1(t) + (\lambda_2)^{\frac{1}{2}}\phi_2(t)\right)\operatorname{sign} a_2$$

$$f_1^*(t) = \frac{(L\lambda_1)^{\frac{1}{2}}}{2^{\frac{1}{2}}|a_2|}\left(a_1 + (a_2^2 - a_1^2)^{\frac{1}{2}}\right)\phi_1(t) + \frac{(L\lambda_2)^{\frac{1}{2}}}{2^{\frac{1}{2}}|a_2|}\left(a_1 - (a_2^2 - a_1^2)^{\frac{1}{2}}\right)\phi_2(t).$$

SOLUTION. By Theorem 2, if we take

$$u(t) = 1/2^{\frac{1}{2}}\left((\lambda_1)^{\frac{1}{2}}\phi_1(t) + (\lambda_2)^{\frac{1}{2}}\phi_2(t)\right)$$

and

$$f_2^*(t) = \operatorname{sign} a_2 L^{\frac{1}{2}} u(t),$$

then $f_1^*(t)$ is any function linearly independent of $f_2^*(t)$ and satisfying $\langle f_1^*, f_1^*\rangle_R = L$ and $\langle f_1^*, u\rangle_R = a_1 L^{\frac{1}{2}}/|a_2|$. Let

$$f_1^*(t) = c_1\phi_1(t) + c_2\phi_2(t);$$

then, from considerations stated above, we get

$$c_1 = (L\lambda_1)^{\frac{1}{2}}\left(a_1 + (a_2{}^2 - a_1{}^2)^{\frac{1}{2}}\right)/\left(2^{\frac{1}{2}}|a_2|\right),$$

$$c_2 = (L\lambda_2)^{\frac{1}{2}}\left(a_1 - (a_2{}^2 - a_1{}^2)^{\frac{1}{2}}\right)/\left(2^{\frac{1}{2}}|a_2|\right).$$

For this solution it can be checked that $\operatorname{tr} aa'M^{-1}(f^*) = a_2^2/L$ as claimed in Theorem 2.

EXAMPLE 2. Here we consider the same model and assumptions as in Example 1, but we want to minimize Var $\hat{\theta}_1$ + Var $\hat{\theta}_2$ in $X = \{h : \|h\|_R^2 \leqslant L\}$. An optimal solution for $\{f_1(t), f_2(t)\}$, by Theorem 1 (ii), is $f_1^{**}(t) = (L\lambda_1)^{\frac{1}{2}}\phi_1(t)$, $f_2^{**}(t) = (L\lambda_2)^{\frac{1}{2}}\phi_2(t)$, $t \in [0, 1]$, which, by the corollary, is also weighted optimal for any symmetric positive definite matrix W. This solution is orthogonal and is different from the $f_k^*(t)$, $k = 1, 2$, in Example 1, which is not orthogonal but linearly independent only. This illustrates that (iii) of Theorem 1 cannot be extended to Theorem 2 for W nonnegative definite matrix.

Acknowledgment. The author wishes to thank Professor G. Wahba for her kind suggestions of this problem during his stay at the University of Wisconsin, Madison. He also appreciates the referee's valuable comments and improvements to an earlier version of this paper.

REFERENCES

1. DOROGOVCEV, A. JA. (1971). Problems of optimal control of a regression experiment. *Selected Transl. Math. Statist. Probability* **10** 35–41.
2. FEDOROV, V. V. (1972). *Theory of Optimal Experiments*. Academic Press, New York.
3. KARLIN, S. and STUDDEN, W. (1966). Optimal experimental designs. *Ann. Math. Statist.* **37** 783–815.
4. KIEFER, J. and WOLFOWITZ, J. (1959). Optimal designs in regression problems. *Ann. Math. Statist.* **30** 271–294.
5. LEVADI, V. S. (1966). Design of input signals for parameter estimation. *IEEE Trans. Automatic Control* **AC-11** 205–211.
6. MEHRA, R. K. (1974). Optimal input signals for parameter estimation in dynamic systems-survey and new results. *IEEE Trans. Automatic Control* **AC-19** 753–768.
7. PARZEN, E. (1959). Statistical inference on time series by Hilbert space methods, I. Technical Report No. 23., Depart. Statist., Stanford Univ.
8. PARZEN, E. (1961). An approach to time series analysis. *Ann. Math. Statist.* **32** 951–989.
9. RIESZ, F. and NAGY, B. SZ. (1956). *Functional Analysis*. Ungar, New York.

DEPARTMENT OF MATHEMATICS
TSING HUA UNIVERSITY
HSINCHU, TAIWAN 300
REPUBLIC OF CHINA