TECHNICAL REPORT NO. 547

January 1979

PARAMETER ESTIMATION IN LINEAR DYNAMIC SYSTEMS

Grace Wahba

Parameter Estimation in Linear Dynamic Systems

by

Grace Wahba

Abstract

We give some fairly general circumstances under which the experimental design problem for linear dynamic systems with correlated noise, can be put in the exact form of the Kiefer-Wolfowitz experimental design problem. Optimal inputs are characterized in terms of their spectral measures with respect to a resolution of the identity in a certain reproducing kernel Hilbert space. Numerical methods established for the Kiefer-Wolfowitz design problem then apply to the numerical determination of an optimal spectral measure.

Keywords: optimal design, dynamic systems

Typist: Mary E. Arthur

Research supported by the U.S. Army Under Contract No. DAAG29-77-G-207.

I. Introduction

Consider the dynamic system

$$y(t) = \left(\sum_{j=1}^{p} \theta_j A_j g\right)(t) + \varepsilon(t), \qquad t \in T, \tag{1}$$

where T is an index set (the unit interval, the integers, the plane, etc.) and $\{\varepsilon(t), t \in T\}$ is a zero mean Gaussian process with $E\varepsilon(s)\varepsilon(t) = K(s,t)$. Letting H_K be the reproducing kernel Hilbert space with reproducing kernel $K(\cdot,\cdot)$, we suppose that the $\{A_j\}$ are a set of bounded linear operators from H_K to H_K which commute with each other and are either all self-adjoint, or are powers of some unitary operator on H_K.

The $\{\theta_j\}$ are to be estimated from the observations $\{y(t), t \in T\}$, when the system is excited by the signal $g \in H_K$. The design problem is to choose g, subject to some restrictions, to maximize some convex, monotone functional $\Phi(\cdot)$ of the inverse M of the covariance matrix $\sum$ of the Gauss Markov estimate $\hat{\underset{\sim}{\theta}}$ of $\underset{\sim}{\theta} = (\theta_1,\ldots,\theta_p)$. A number of special cases of this problem have appeared in the literature, some of practical importance. See Box and Jenkins [2], Viort [13,14], Mehra [8], Payne, Goodwin and Zarrop [10], Goodwin and Payne [6], Keviczky [7], Dogorovcev [4]. Spruill and Studden [11,12], and Chang [3] discuss closely related problems.

Behind some of the known results for special cases of this problem there is, either explicitly or implicitly a reduction of the design problem for (1) to the so called Kiefer-Wolfowitz design problem. In this latter problem the model is

$$y(x) = \sum_{j=1}^{p} \theta_j f_j(x) + \varepsilon(x), \qquad x \in [-1,1]$$

where observations on $\varepsilon(x)$, $x \in [-1,1]$ are i.i.d., $N(0,\sigma^2)$. One chooses $x_1,\ldots,x_N \in [-1,1]$, (possibly coincident) to maximize $\Phi(M)$, where M is the inverse of the covariance matrix $\sum$ of the Gauss-Markov estimate of $\underset{\sim}{\theta}$ based on observations $\{y(x_i), i = 1,2,\ldots,N\}$. The jk^{th} entry of M is given by

$$M_{jk} = \Sigma_{jk}^{-1} = \int_{-1}^{1} f_j(x)f_k(x)d\xi(x) \tag{2}$$

where ξ is the design measure (=fraction of $x_i \leq x$).

Given the $\{f_j\}$, the design problem is thus that of choosing ξ in (2) to minimize $\Phi(M)$. An optimal ξ can always be found as a jump function with at most $\frac{1}{2}p\,(p+1)+1$ jumps. The literature on this problem is extensive. For background see the book of Fedorov [5]. There is a developing literature on numerical calculation of an optimal ξ given $f_1,\ldots,f_p$, we mention only the recent work of Wu [14,15].

The purpose of this note is to establish some rather general conditions under which the design problem for (1) can be reduced to the form of the design problem for (2). We make this reduction in Section II and give some examples in Section III.

II. Reduction of the Linear Dynamic System Design Problem to the Kiefer-Wolfowitz Design Problem

By Parzen [8], we have that the covariance matrix $\sum$ for the Gauss-Markov estimate of $\underset{\sim}{\theta}$ in the system (1) is given by

$$\Sigma_{jk}^{-1} = \langle A_j g, A_k g\rangle_K \equiv M_{jk}$$

where $\langle\cdot,\cdot\rangle_K$ is the inner product in H_K.

Let A_0 be an arbitrary self adjoint operator on H_K and let $A_1,\ldots,A_p$ be p[illegible]ded operators which commute with every bounded

operator which commutes with A_0. Then there exists a resolution of the identity in H_K, (see the Appendix for the definition), call it E_λ, such that

$$A_0 g = \int_{-\infty}^{\infty} \lambda\, dE_\lambda g \qquad \text{all } g \in \text{domain of } A_0$$

and there exist ([1], p.38, Theorem 2), p bounded functions $f_1,\ldots,f_p$ such that

$$A_j g = \int_{-\infty}^{\infty} f_j(\lambda) dE_\lambda g \qquad j = 1,2,\ldots,p\,, \qquad g \in H_K\,,$$

([1],p.77-79). Then by the properties of E_λ,

$$\langle A_j g, A_k g\rangle_K = \int_{-\infty}^{\infty} f_j(\lambda)f_k(\lambda)d\xi(\lambda) \equiv M_{jk} \tag{3}$$

where $\xi(\lambda)$ is the measure

$$\xi(\lambda) = ||E_\lambda g||_K^2\,.$$

Next, let $A_1,\ldots,A_p$ be powers of some unitary operator U, say $A_k = U^k$. Then there exists ([1],p.21) a resolution of the identity E_λ such that

$$A_k g \equiv U^k g = \int_0^{2\pi} e^{ik\lambda} dE_\lambda g\,,$$

and, similarly

$$\langle A_j g, A_k g\rangle_K = \int_0^{2\pi} e^{i(j-k)\lambda} d\xi(\lambda) \equiv M_{jk} \tag{4}$$

where

$$\xi(\lambda) = ||E_\lambda g||_K^2\,.$$

Since $\xi(\infty) = ||g||_K^2$, if the only constraint on g is $||g||_K^2 \leq C$, then we see that the problem of finding the spectral distribution ξ of an optimal g in (3) and (4) is exactly the same as finding an optimal design measure in (2). Note that knowledge of $\xi(\lambda)$ determines g up to the norm of its projections on an increasing sequence of orthogonal subspaces whose limit is H_K. Two g's with the same ξ are of equal merit with respect to any design criteria which is a function only of M.

Constraints on g other than $||g||_K^2 \leq C$ are of some interest. Suppose that the constraint is $||Bg||_K^2 \leq C$, where B is a self adjoint operator in H_K commuting with each E_λ. Then B has a representation

$$Bg = \int b(\lambda)dE_\lambda g, \qquad g \in H_K . \tag{5}$$

Define

$$\beta(\lambda) = ||E_\lambda Bg||_K^2 . \tag{6}$$

Then

$$d\beta(\lambda) = b^2(\lambda)||dE_\lambda g||_K^2$$

and

$$<A_j g, A_k g>_K = \int \frac{f_j(\lambda)f_k(\lambda)}{b^2(\lambda)} d\beta(\lambda) . \tag{7}$$

Since $\beta(\infty) = ||Bg||_K^2$, the design problem with the constraint $||Bg||_K^2 \leq C$ reduces to finding an optimal $d\beta$, equivalently it reduces to a problem of the form of (2), with $f_j(\lambda)$ replaced by $f_j(\lambda)/b(\lambda)$.

In particular, let K be the integral operator defined on H_K by

$$(Kg)(t) = \int_T K(t,s)g(s)ds, \qquad t \in T \tag{8}$$

and $K^{1/2}$ its symmetric square root. (We are assuming at least that K is densely defined on $L_2(T)$). Then $||K^{1/2}g||_K^2 = ||g||_{L_2}^2 = \int g^2(t)dt$ by the properties of reproducing kernel inner products. If K commutes with each E_λ, then we may set $B = K^{1/2}$ in the above and have $\beta(\infty) = \int_T g^2(t)dt$. Thus the constraint $\int_T g^2(t)dt \leq C$ also leads to a problem of the form of (2).

For other constraints of the form $g \in C$ where C is any closed bounded convex set in H_K, it is useful to know that the functional $\tilde{\Phi}(g)$ defined on H_K by $\tilde{\Phi}(g) = \Phi(M)$ is convex. Then we know that maximizer(s) of $\tilde{\Phi}(\cdot)$ are found on the boundary of C. Φ convex and monotone implies $\tilde{\Phi}$ is convex as follows: (By Φ monotone is meant M_2-M_1 non-negative definite entails $\Phi(M_2) \geq \Phi(M_1)$.)

i) $||dE_\lambda g_1||_K^2 \leq ||dE_\lambda g_2||_K^2 \Rightarrow M(g_2)-M(g_1)$

is non-negative definite, and then $\tilde{\Phi}(g_1) \leq \tilde{\Phi}(g_2)$.

ii) $$||dE_\lambda(\alpha g_1+(1-\alpha)g_2||_K^2 \leq \alpha^2||dE_\lambda g_1||_K^2$$
$$+2\alpha(1-\alpha)||dE_\lambda g_1||_K||dE_\lambda g_2||_K+\alpha^2||dE_\lambda g_2||_K^2$$
$$\leq \alpha||dE_\lambda g_1||_K^2+(1-\alpha)||dE_\lambda g_2||_K^2$$

iii) Thus $$\tilde{\Phi}(\alpha g_1+(1-\alpha)g_2) \leq \Phi(\alpha M(g_1)+(1-\alpha)M(g_2))$$
$$\leq \alpha\tilde{\Phi}(g_1)+(1-\alpha)\tilde{\Phi}(g_2) .$$

The set $C = \{g: \ |g(t)|^2 \leq C, \ t \in T\}$ is closed and convex in any reproducing kernel space. The boundary of C consists of g for which $|g(t)|^2 \leq C$ with equality for some t. Bang-bang functions (i.e. functions for which $g(t) = \pm C$ for all t with both "+" and "-" obtaining) will not, however, be in H_K if H_K is a space of continuous functions.

III. Examples

We illustrate these results with several examples. First suppose K defined in (8) considered as an operator on $L_2(T)$ is compact, that is,

K possess an (L_2-orthonormal) countable system of eigenfunctions $\{\phi_\nu\}$ with associated eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq 0$ with $\lambda_\nu \to 0$. (If $\sum \lambda_\nu^2 < \infty$ then $K(s,t) = \sum_\nu \lambda_\nu \phi_\nu(s)\phi_\nu(t)$). For concreteness, assume 0 is not an eigenvalue of K, then the $\{\phi_\nu\}$ are complete. Let the $\{A_j\}$ be defined by

$$A_j \phi_\nu = a_{j\nu}\phi_\nu \qquad \nu = 1,2,\ldots$$
$$j = 1,2,\ldots,p ,$$

with the $\{a_{j\nu}\}$ uniformly bounded. Letting $A_0 = K$, the A_j commute with every operator that commutes with A_0, and

$$E_\lambda g = \sum_{\nu:\lambda \leq \lambda_\nu} g_\nu \phi_\nu$$

where

$$g_\nu = \int_T \phi_\nu(t) g(t)dt ,$$

and

$$\xi(\lambda) \equiv ||E_\lambda g||_K^2 = \sum_{\nu:\lambda_\nu \leq \lambda} \frac{g_\nu^2}{\lambda_\nu} .$$

The f_j need only be defined at the possible jumps of ξ, and are given by

$$f_j(\lambda) = a_{j\nu} , \qquad \lambda = \lambda_\nu .$$

If $B = K^{1/2}$, then b and β defined in (5) and (6) are given by

$$b(\lambda) = \lambda^{1/2}, \ \lambda = \lambda_1, \lambda_2, \ldots$$

$$\beta(\lambda) = \sum_{\nu:\lambda_\nu \leq \lambda} g_\nu^2$$

An optimal g subject to $||g||_K^2 \leq C$, or $||g||_{L_2}^2 \leq C$ can be found as a linear combination of at most $\ell = \frac{1}{2}p(p+1)+1$ eigenfunctions of K, that is, an optimal g is of the form

$$g = \sum_{j=1}^{\ell} c_j \phi_{\nu_j}$$

for some $c_1,\ldots,c_\ell$, $\nu_{j_1},\ldots,\nu_{j_\ell}$.

Next, let $T = \ldots -1,0,1,\ldots$ and let

$$K(s,t) = \int_0^{2\pi} e^{i\nu(s-t)}\lambda(\nu)d\nu$$

where, for simplicity we assume $0 < \lambda_{min} \leq \lambda(\nu) \leq \lambda_{max} < \infty$. $\varepsilon(t)$ is a stationary time series with spectral density $\lambda(\nu)$, $0 \leq \nu \leq 2\pi$.

Functions g with a representation

$$g(t) = \int_0^{2\pi} e^{it\nu}\tilde{g}(\nu)d\nu$$

are in H_K if and only if

$$\int_0^{2\pi} \frac{|\tilde{g}(\nu)|^2}{\lambda(\nu)} d\nu \equiv ||g||_K^2 < \infty .$$

Let U^k, $k = 1,2,\ldots$ be the shift operators defined on H_K by

$$(U^k g)(t) = g(t+k) = \int_0^{2\pi} e^{ik\nu} e^{it\nu}\tilde{g}(\nu)d\nu , \qquad t = \ldots -1,0,1,\ldots$$

We have

$$(E_\lambda g)(t) = \int_0^{\lambda} e^{it\nu}\tilde{g}(\nu)d\nu$$

and

$$\xi(\lambda) = ||E_\lambda g||_K^2 = \int_0^{\lambda} \frac{|\tilde{g}(\nu)|^2}{\lambda(\nu)} d\nu .$$

If Bg is defined by

$$(Bg)(t) = \int_0^{2\pi} e^{it\nu} b(\nu)\tilde{g}(\nu)d\nu$$

and $\beta(\lambda) = ||E_\lambda Bg||_K^2$, then

$$d\beta(\lambda) = b^2(\lambda)d\xi(\lambda) .$$

An optimum ξ subject to the constraint $||g||_K^2 \equiv \int_0^{2\pi} \frac{|\tilde{g}(\nu)|^2}{\lambda(\nu)} d\nu \equiv \xi(2\pi) \leq C$, or the constraint $||Bg||_K^2 = \int_0^{2\pi} \frac{|b(\nu)|^2}{\lambda(\nu)} |\tilde{g}(\nu)|^2 d\nu = \beta(2\pi) \leq C$ can be found as a jump function with at most $\ell = \frac{1}{2}p\,(p+1)+1$ jumps, at $\nu = \nu_1, \nu_2, \ldots, \nu_\ell$, say. The associated g's are of the form

$$g(t) = \sum_{j=1}^{\ell} c_j e^{it\nu_j} .$$

This was one of the cases considered by Viort [13,14], see also Mehra [8].

The boundedness assumption on the A_j is unduly restrictive since useful examples can be found where the results hold for unbounded operators. Let $T = [0,\tau]$ and let

$$K(s,t) = \sum_{\nu=-\infty}^{\infty} \lambda_\nu e^{2\pi i\nu(s-t)/\tau}, \qquad \lambda_\nu = \lambda_{-\nu} .$$

If $\lambda_\nu = 2\pi\lambda(2\pi\nu/\tau)/\tau$ where $\lambda(\omega)$ is a "nice" integrable spectral density on $(-\infty,\infty)$, then $K(s,t)$ is a circulant approximation to the covariance on $[0,\tau]$ of a stationary stochastic process with covariance $\tilde{K}(s,t)$ given by

$$\tilde{K}(s,t) = \int_{-\infty}^{\infty} \lambda(\omega)e^{i\omega(s-t)}d\omega.$$

Then if

$$Kg = \int \lambda \, dE_\lambda \, g$$

we have

$$A_j g \equiv g^{(j-1)} = \int f_j(\lambda)dE_\lambda g$$

with $f_j(\lambda) = (2\pi\nu/\tau)^{j-1}$, $\lambda = \lambda_\nu$. (Obvious modification is made if the λ_ν's are not distinct.) Common examples of $\tilde{\Phi}(\cdot)$ will be unbounded on bounded sets in H_K. The design problem will make sense if the constraints require g to be in a set on which $\tilde{\Phi}$ is bounded, for example

$$\{g: \int_T (g^{(p)})^2 dt \leq C.\}$$

Here, we can let $Bg = g^{(p)}$, and proceed as before:

If $\beta(\lambda) = ||E_\lambda Bg||_K^2$, then

$$d\beta(\lambda) = b^2(\lambda)d\xi(\lambda)$$

with $b^2(\lambda) = (2\pi\nu/\tau)^p$, $\lambda = \lambda_\nu$.

Appendix

Definition: ([1], p.16.). A resolution of the identity is a one parameter family of projection operators E_λ, where λ runs through a finite or infinite interval $[\alpha,\beta]$, which satisfies the following conditions:

a) $E_\alpha = 0$, $E_\beta = E$ (the identity)

b) $E_{\lambda-0} = E_\lambda$ $(\alpha < \lambda < \beta)$

c) $E_\mu E_\nu = E_{\min(\mu,\nu)}$

If the interval $[\alpha,\beta]$ is infinite, then $E_{-\infty} = \lim_{\lambda\to-\infty} E_\lambda$, $E_\infty = \lim_{\lambda\to\infty} E_\lambda$.

REFERENCES

[1] Akhiezer, N.I., and Glasman, I.M., Theory of Linear Operators in Hilbert Space, Vol. II, Translated from the Russian by Merlind Nestell. Ungar, (1963).

[2] Box, G.E.P., and Jenkins, G., Time Series Analysis, forecasting and control, Holden Day (1970).

[3] Chang, D.S., Design of optimal control for a regression problem, to appear, Ann. Statist.

[4] Dogorovcev, A.J., Problems of optimal control of a regression experiment. Selected Translations in Math. Stat. and Prob., 10, 35-41, American Math Society (1971).

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

[6] Goodwin, G.C., and Payne, R.L., Dynamic System Identification: Experiment Design and Data Analysis. New York: Academic Press (1977).

[7] Keviczky, L., "Design of Experiments" for the Identification of Linear Dynamic Systems Technometrics, 17, 3(1975), 303-307.

[8] Mehra, R.K., Optimal input signals for parameter estimation in dynamic systems - survey and new results. IEEE Trans. Automat. Contr. AC-19, 753-768 (1974).

[9] Parzen, E., An approach to time series analysis. Ann. Math. Statist. 32, 951-989 (1961).

[10] Payne, R.L., Goodwin, G.C., and Zarrop, M., Frequency domain approach for designing sampling rates for system identification. Automatica, 11, 189-191 (1975).

[11] Spruill, C., and Studden, W.J., Optimum designs when the observations are second order processes. J. Multivar. Anal. 8, 153-172 (1978).

[12] Spruill, C., and Studden, W.J., A Kiefer-Wolfowitz Theorem in a stochastic process setting, manuscript (1978).

[13] Viort, B., D-optimal designs for dynamic models: Part I, Theory. Tech. Report 314, Department of Statistics, University of Wisconsin, Madison (1972).

[14] Viort, B., D-optimal designs for dynamic models: Part II, applications. Tech. Report 316, Department of Statistics, University of Wisconsin, Madison (1972).

[15] Wu, Chien-Fu., Some iterative procedures for generating non-singular optimal designs. Commu. Stat., Theory and Methods, A7 14(1978), 1399-1412.

[16] Wu, Chien-Fu., Some algorithmic aspects of the theory of optimal designs. Ann. Statist. 6,6, 1286-1301 (1978).

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TR#547	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) PARAMETER ESTIMATION IN LINEAR DYNAMIC SYSTEMS		5. TYPE OF REPORT & PERIOD COVERED 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Grace Wahba		8. CONTRACT OR GRANT NUMBER(s) DAAG29-77-G-207
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics University of Wisconsin Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Research Office P.O. Box 12211 Research Triangle Park, N.C. 27709		12. REPORT DATE January 1979 13. NUMBER OF PAGES 11
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this document is unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

optimal design, dynamic systems

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

We give some fairly general circumstances under which the experimental design problem for linear dynamic systems with correlated noise, can be put in the exact form of the Kiefer-Wolfowitz experimental design problem. Optimal inputs are characterized in terms of their spectral measures with respect to a resolution of the identity in a certain reproducing kernel Hilbert space. Numerical methods established for the Kiefer-Wolfowitz design problem then apply to the numerical determination of an optimal spectral measure.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)