1982

D-optimal input design for parameter estimation of linear and distributed parameter systems

Zahid H. Qureshi
University of Wollongong

D-OPTIMAL INPUT DESIGN FOR PARAMETER ESTIMATION OF LINEAR AND DISTRIBUTED PARAMETER SYSTEMS

A thesis submitted in fulfilment of
the requirements for the award of
the degree of

DOCTOR OF PHILOSOPHY

from

THE UNIVERSITY OF WOLLONGONG

by

ZAHID H. QURESHI, B.Sc.Eng.,

Department of Electrical
and Computer Engineering

September 1982

Dedicated to my parents

This is to certify that the work reported in this thesis was done by the author, unless specified otherwise, and that no part of it has been submitted in a thesis to any other University.

Zahid H Qureshi
....

Zahid H. Qureshi

ABSTRACT

This thesis is concerned with the optimal input design problem for parameter estimation of linear, single input single output, discrete-time dynamic systems. Some aspects of experimental design for parameter estimation of distributed parameter systems are also studied.

The experiment design problem is to choose the experimental conditions subject to constraints such that the information from an experiment is maximized in some sense. The optimality criterion employed in this thesis is the determinant of Fisher information matrix (D-optimality).

The design of optimal inputs for parameter estimation of linear dynamic systems comprises the major contribution of this thesis. The open-loop input design is studies for systems subject to input and output power constaints. A method is given to obtain an optimal input design for an autoregressive model with output power constraints. It is shown that the optimal input frequencies can be obtained by solving a set of non-linear equations without recourse to optimization techniques involving the calculation of determinants. The problem of optimal input design for estimating part of the system parameters of a general model is investigated. Some results from D-optimal designs are extended for the D_s-optimal case. A comparison of minimal uniform input designs and a D-optimal design is also carried out, and

the D-efficiency of minimal uniform input designs is illustrated by considering first order systems.

The relationship between the experimental conditions and the achievable accuracy in parameter estimation for distributed parameter systems is also investigated. In distributed parameter systems, besides the boundary perturbations, another important design variable is available, namely, the spatial location of measurement sensors. A method to design optimal experiments for parameter estimation of a general distributed parameter system is proposed and illustrated by typical designs for a parabolic and a hyperbolic system.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. T.S. Ng, for his guidance and encouragement throughout this research.

My special thanks go to Andreas Moqhali and Sarath Perera for many useful suggestions and discussions which helped towards the completion of this thesis. I am grateful to the staff of the Department of Electrical Engineering at the University of Wollongong for their assistance, particularly Professor B.H. Smith and Dr. K.J. McLean.

I would also like to acknowledge the assistance offered by the Department of Mathematics at the University of Wollongong. In particular, I am grateful to Dr. Barry Quinn for the many stimulating discussions and helpful comments.

My thanks also to the University of Wollongong for the financial support in the form of a Postgraduate Research Scholarship. My sincere thanks to Enid Sherwin for her keen interest in typing the thesis.

Finally, I am deeply grateful to my mother for her constant encouragement from afar. I also wish to express my sincere thanks to my wife, Uzma, for her patience and understanding during this research.

LIST OF SYMBOLS

a_i	coefficient of polynomial A
$A(\cdot)$	denominator polynomial of system transfer function (t.f.)
b_i	coefficient of polynomial B
$B(\cdot)$	numerator polynomial of system t.f.
c_i	coefficient of polynomial C
$C(\cdot)$	denominator polynomial of noise t.f.
d_i	coefficient of polynomial D
$d(\cdot,\cdot)$	generalized variance
$D(\cdot)$	numerator polynomial of noise t.f.
e_i	white noise sequence
E	expectation operator
$h(\cdot)$	$u \to \partial\varepsilon/\partial\theta$
k	discrete time
L	log likelihood function
m	degree of $B(\cdot)$
M	information matrix
$\mathcal{M}$	set of information matrices
n	degree of $A(\cdot)$
N	number of data points
p	number of system parameters
q	degree of $C(\cdot)$
r	degree of $D(\cdot)$
t	continuous time
u	system input
y	system output
z^{-1}	unit backward shift operator

δ	Dirac delta function
δ_{ij}	Kronecker delta
ε_i	residual sequence
θ	parameter vector
λ_i	power proportion
ξ	design measure
Ξ	set of normalized design measures
σ^2	variance
χ	experimental region
ω	frequency

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

CHAPTER 1: INTRODUCTION

1.1 INTRODUCTION

This thesis is concerned with the optimal input design problem for parameter estimation of linear dynamic systems. Some aspects of experimental design for parameter estimation of distributed systems are also studied.

Experiment design plays an important role in the fields of science and technology where complicated and expensive experiments are carried out. The problem of extracting maximal information with finite resources leads to the planning of experiments in an optimal manner.

The philosophy of experimental design is that one has freedom, to some extent, to decide on some relevant environmental conditions under which the process is operating. For dynamic system identification, some of

the variables which can be manipulated for optimal experimental design are the choice of input and measurement ports, choice of test signals, sampling instants, pre-sampling filters and feedback settings. Each of these variables has a significant bearing upon the information return of the experiment and usually their effects are inter-related.

While designing experiments, one must also consider the constraints which are imposed on the experimental conditions. In fact, the real purpose of experiment design is to maximize the information content of the data within the limits imposed by the given constraints. Typical constraints are amplitude and power constraints on the inputs and outputs, total time allotted to an experiment, total number of samples that can be taken and the sampling rate.

In most real world systems, a lumped representation of a process (described by ordinary differential equations/difference equations) is often adequate to describe the system behaviour. However, in many applications which involve atmospheric and other natural phenomena, and in many basic industries such as mineral processing, steel, petroleum, chemical, glass and nuclear power, a distributed parameter representation of the process is required in order to describe the process more accurately and to implement more effective control strategies. Due to this wide interest in distributed

systems, the experimental design problem for systems described by partial differential equations is also studied.

1.2 IDENTIFICATION AND EXPERIMENTAL DESIGN

The identification problem is frequently formulated as follows: Given a class of models, a criterion and measurements of input and output signals, find the particular model which fits the experimental data best in the sense of the given criterion.

The selection of the class of models, the class of input signals and the criterion is mainly based on both the *a priori* knowledge and identification purpose. The four steps in system identification are shown schematically in Figure 1.1.

Throughout this thesis, we consider systems with known model structures and the goal of identification is to estimate the process parameters.

The dependence of the optimal experimental conditions on the unknown parameters is a common occurrence in optimal experiment design for dynamic systems. To overcome the difficulty, two approaches have been used. The first approach uses the Bayesian formulation by assuming that an *a priori* probability distribution for the parameter vector is given. The second approach assumes that a nominal value of the parameter vector is given. In this thesis we have adopted the latter viewpoint.

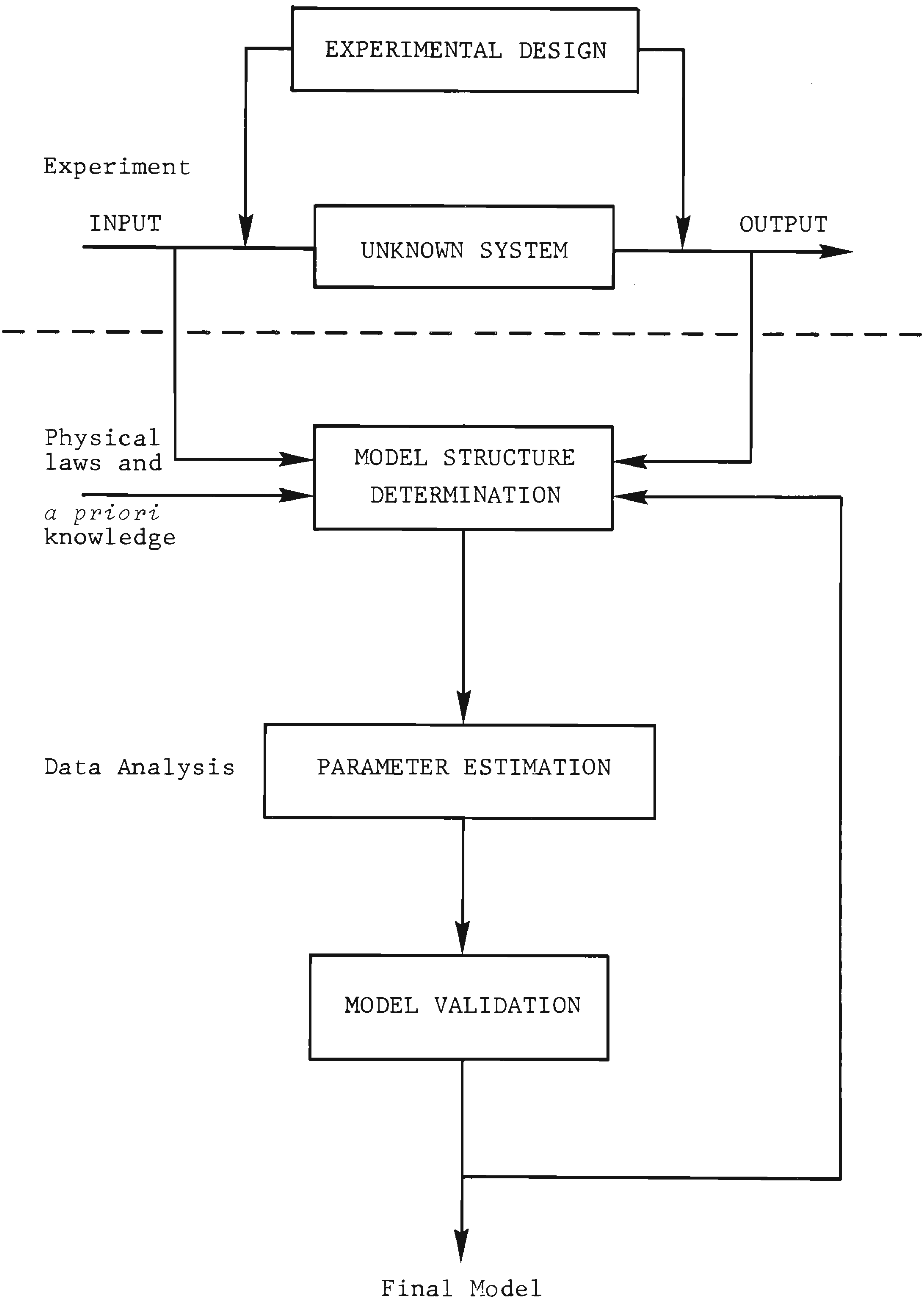


Fig. 1.1: Steps in system identification

Figure 1.1 shows the location of experimental design and parameter estimation in the overall identification procedure. The study of optimal experiments for parameter estimation is of interest in a number of areas of research, e.g. heavy industry, process control, chemical engineering, etc., where the final goal of identification is the determination of specific parameter values.

The significance of experiment design for parameter estimation is reflected in a wide spectrum of publications in both control and statistical literature. In the control literature, interest is mostly on dynamic systems. Aström and Eykhoff [A3] and Nieman et al. [N7] give a comprehensive survey of parameter estimation techniques and problems associated with process identification. Gustavsson et al. [G9] review a number of aspects of parameter estimation for closed loop systems. Applications of identification techniques to aeronautics has been reviewed by Rault [R1], while those for physical processes have been reviewed by Gustavsson [G8]. Several books on system identification are also available, e.g. Eykhoff [E1], Sage and Melsa [S1], Graupe [G7], Mehra and Lainiotis [M5], Goodwin and Payne [G6], and Zarrop [Z3].

In the next Section (Sec.1.3) a literature survey on input design problem is present while in Section 1.4, a discussion on distributed parameter systems is presented.

1.3 OPTIMAL INPUT DESIGN

STATISTICAL BACKGROUND

The basic theory of optimal experiments was developed for linear regression models with independent errors. The interested reader is directed to Kiefer [K1] and Kiefer and Wolfowitz [K2] for historical background. The book by Fedorov [F1] and the survey paper by St. John and Draper [S8] provide a comprehensive discussion and survey of the D-optimal design problems for the regression models described by (1.3.1):

$$y_j = \theta^T f(x_j) + e_j \quad j=1,2,\ldots, N \tag{1.3.1}$$

where:

y_j is the jth observation, and

$\{e_j\}$ is a sequence of uncorrelated and identically distributed random variables with zero mean and variance σ^2,

the p x 1 vector $f(\cdot)$ is assumed known and continuous on some compact set X (experimental region),

and the observations y_i and y_k are uncorrelated.

From the observations $\{y_j\}$ a minimum variance unbiased linear estimator $\hat{\theta}$ of the p x 1 vector θ can be derived by the least-squares procedure. The covariance matrix of $\hat{\theta}$ is given by:

$$\operatorname{cov} \hat{\theta} = \sigma^2 \left[\sum_{j=1}^{N} f(x_j) f^T(x_j) \right]^{-1}$$

The design problem consists of selecting vectors x_i, $i=1,\ldots,N$ from X such that the design defined by these N vectors is, in some specified sense, optimal. This is usually carried out by choosing the design to minimize some scalar function of $\operatorname{cov} \hat{\theta}$.

Concept of Design Measure:

Kiefer and Wolfowitz [K2] extended this concept of design by introducing a measure ξ on X. The problem then is to choose an optimal design measure which specifies the proportion of observations to be made at different points in X. Kiefer and Wolfowitz [K3] then showed that the two different design criteria, D-optimality and G-optimality (min-max design), are equivalent.

DYNAMIC SYSTEMS

Under certain conditions, the concept of optimal experimental design problem for regression models can be extended to dynamic systems (see Viort [V2], Mehra [M4]), and is referred to as the input synthesis, or the optimal input design problem.

It has long been recognized that the choice of input signal in an identification experiment has a significant bearing upon the parameter estimation accuracy. This has spurred increasing interest in the design of inputs that are optimal in the sense that they allow maximum information to be extracted from the identification experiment. For a comprehensive survey of literature on input design, see Mehra [M4].

Some Early Papers:

The optimal input design problem has been studied by a number of authors, using both the time domain and the frequency domain approach.

Levin [L3] appears to be one of the first authors to have considered an optimal input design problem for dynamic systems. He discussed the problem of estimating the impulse response of a discrete-time, single-input single-output (SISO) linear system in the presence of white noise. He showed that the optimal energy constrained input for minimum variance estimation is a white noise sequence.

Another early paper on optimal input for identification was by Litman and Huggins [L4] in 1963. In their work, an energy constraint was placed on the input and the determinant of the covariance of a two parameter system was minimized.

In 1966, Levadi [L2] considered a linear, time varying, moving-average model with non-stationary coloured output noise. An energy constraint was placed on the input signal and the necessary conditions for optimality were derived.

Control Theoretic Formulation:

If an asymptotically efficient unbiased parameter estimator (e.g., maximum likelihood) is used, then for long data lengths the parameter covariance matrix can be approximated by the inverse of Fisher information matrix (Cramer-Rao lower bound). The input design problem can then be formulated in control-theoretic terms with the input chosen to optimize some suitable scalar function of the information matrix. Nahi and Wallis [N3], Aoki and Staley [A1], Nahi and Napjus [N2] and Mehra [M1] used the trace of a weighted information matrix, which leads to a quadratic problem with analytic solution. However, this criterion can result in inputs which lead to a singular information matrix and are unsuitable for parameter identification purposes. This has been pointed out by Goodwin [G2] and Zarrop and Goodwin [Z1].

Input Amplitude Constraint:

Keviczky and Banyasz [K5] considered a linear, discrete-time dynamic system and placed an amplitude constraint on the input. The criterion employed was

the minimization of the determinant of the inverse information matrix. They found that the optimal amplitude constrained input is of the bang-bang type, and also specified the instants when the input switches polarity.

FREQUENCY DOMAIN APPROACH

The study of asymptotic properties of the information matrix leads to a frequency domain formulation of the input design problem. This approach has led to many simplifications and greater insight into the design problem.

Goodwin and Payne [G5] and Van den Bos [V1] have shown that, for large record lengths, the optimal input signal for linear SISO system parameter estimation can be characterized by its spectral properties.

Normalized Input Designs:

A natural constraint to impose on the input is that its power be restricted and this leads to the consideration of normalized designs, i.e. the total power in the input spectrum be restricted to unity. The design space in the frequency domain approach is the finite interval $(-\pi,\pi)$, and the design measure defines the distribution of the total power over the range of frequencies.

Discrete Input Designs:

Mehra [M2][M3][M4] has shown that, for any normalized input design with a mixed (continuous and discrete) spectrum, another input design can be found with a purely discrete spectrum containing no more than $[p(p+1)/2] + 1$ frequencies, where p is the number of parameters to be estimated, such that they have the same information matrix. Thus the input design problem has been reduced to a finite dimensional optimization problem. Mehra also extended the equivalence theorem [K3] for dynamic systems.

Payne and Goodwin [P2] have shown that, for SISO systems, the maximum number of frequencies required in the input spectrum is less than or equal to p. This simplification results from the special structure of the information matrix. This result leads to a significant reduction in the number of independent variables from $p(p+1)$ to $2p$ required for optimality.

Recently Zarrop [Z2] used a geometric approach, based on classical Tchebycheff system theory, to analyze the input design problem for parameter estimation. He considered a linear, SISO, discrete-time general model, and showed that under certain system order constraints, a D-optimal design could be achieved for both the input and output power constraint cases by $p/2$ or $p+1/2$ input frequencies.

If a D-optimal design exists containing p/2 or p+1/2 frequencies, then the power proportions have the values 2/p if $\omega_i \in (0,\pi)$ and 1/p if $\omega_i \in 0$ or π([G6][Z3]).

Output Power Constraint:

In many situations a more natural constraint is on the system output fluctuations. For example, in an industrial process, the quality of product output usually has to be regulated within certain prescribed tolerance limits; alternatively, in a biological experiment or a clinical trial, the well-being of the object under treatment may be of prime importance.

Constrained output design has been considered by a number of authors. Söderström, Ljung and Gustavsson [S4][S5] and Söderström [S6] have examined a first order system with constrained output variance. This study has raised a number of interesting conjectures regarding the achievable accuracy in closed loop experiments. Various rules of thumb have also been suggested for more general models. The autoregressive model has been studied by Ng, Goodwin and Payne [N4]. It was shown that a minimum covariance control law, together with white perturbation signals, achieves D-optimality. Extension of the above result to the general model has also been attempted by Ng, Goodwin and Söderström [N5]. It was shown that the minimum variance control strategy gives a D-optimal experiment for a new set of parameters which are related

via a simple transformation to the original parameters. However, Ng et al. [N5] could not show that the design is also D-optimal for the original parameters.

SEQUENTIAL DESIGNS

In many experiments, the experimental conditions change during the time they are conducted. Under such conditions it is not suitable to construct a design which would specify the distribution of all resources allotted for the given investigations (statistical design). Such a design would usually be non-optimal, since it is impossible beforehand to predict all possible situations, e.g. further information obtained during the course of investigation, small variations in the parameter values.

In such cases, it is more expedient to turn to sequential methods of designing experiments, in the determination and refinement of unknown parameters. The idea of a sequential design consists of the following: the resources (e.g. time) are divided into small portions; the experiment is divided into several steps and at each step a design is carried out, using one portion of the resources; an analysis of the experiment is conducted after each step; the experiment ceases as soon as a given characteristic or precision of the estimates of the parameters attains a prescribed value.

Sequential Design Algorithms:

Mehra [M2] proposed a sequential design algorithm for dynamic systems based on that of Fedorov [F1], converging globally to a D-optimal design. This algorithm may result in a large number of frequencies in the final design and the rate of convergence can be slow.

Zarrop [Z3] proposed variations in the sequential design algorithms to improve the rate of convergence and to reduce the number of frequencies in the design. He introduced the so-called S-algorithms which are globally convergent. He proposed a modification to the algorithm by Mehra [M2], and extended Atwood's algorithm [A5] to remove power from certain frequencies in the input spectrum. He also employed a frequency rounding-off procedure in his algorithms, and carried out comparisons of several algorithms, using first and second order models. From this comparison, he found that, for $p > 2$, the algorithm which allows the removal of frequencies gives substantial improvement in the rate of convergence and results in final designs containing only p frequencies.

THE SHAPING FILTER APPROACH

Another approach to synthesize optimum inputs is by passing white noise through a shaping filter. The design problem now is to estimate the filter parameters

for optimal estimation of system parameters. Upadhyaya and Sorenson [U1] considered a general ARMA filter, while Berger [B1] constrained the filter to be a MA filter. This approach appears quite attractive from a computational point of view. However, the results derived are at best sub-optimal as the order of the filter used is prefixed. More recently, Stoica and Söderström [S7] showed that an optimal input can be synthesized by white noise passing through an ARMA filter, and they also obtained the required filter order for optimality.

NON-LINEAR SYSTEMS

Optimal input design for non-linear dynamic systems has also been studied by some authors:

Goodwin [G3] described a design procedure to synthesize an optimal input for minimum variance estimation of the parameter and states in models of non-linear dynamic systems;

Mehra [M4] discussed the time-domain input design for non-linear and distributed parameter systems without process noise. The main difficulty occurred in calculating the information matrix, and in optimizing a scalar objective function of the information matrix with respect to the input.

1.4 IDENTIFICATION OF DISTRIBUTED PARAMETER SYSTEMS

Basically the main theoretical difficulty for identifying systems described by partial differential equations is due to the infinite dimensionality of the state space. Two approaches are normally used to face this problem:

(i) approximation of the infinite-dimensional model by a finite-dimensional model (described by ordinary differential equations, difference equations and algebraic equations);

(ii) application of optimization techniques directly to the infinite-dimensional model.

For a detailed literature review of different approximation techniques and parameter estimation methods, see the survey paper by Kubrusly [K6].

In distributed parameter systems, another important design variable is available, namely, the spatial location of measurement sensors. Besides the usual constraints met in dynamic systems, constraints in distributed systems include the allowable boundary perturbations, the initial conditions, the nature of measurement sensors, the number of sensors, the allowable measurement zone, the total number of samples, the total experimental time, and the allowable response amplitude.

Parameter identification problems for distributed systems have been studied by a number of authors. Perdreauville and Goodson [P3] used integration by parts together with measurement data to yield a set of algebraic equations; Carpenter et al. [C1] used the method of characteristics and stochastic approximation to yield parameter estimates for hyperbolic systems; Seinfeld [S2] and Tzafestas and Nightingale [T2] used non-linear filters; Polis et al. [P4] used Galerkin's method to transform the partial differential equation into a set of ordinary differential equations and obtained parameter estimates by various optimization techniques.

In the above studies, parameter estimates were obtained for given experimental conditions. The relationship between information content of measurement data and experimental conditions has not, as yet, been studied.

A closely related problem is the optimal location of measurement sensors for state estimation in distributed systems. This problem has been examined (see, for example, Yu & Sienfeld [Y1] and Goodson & Klein [G1]). However, the extension to parameter estimation is not straightforward and has not been pursued.

1.5 THESIS REVIEW

This thesis is mainly concerned with the optimal input design for parameter estimation of linear dynamic, SISO, systems. The open-loop design is considered in this work. Optimal experimental designs are also studied for identifying the parameters of distributed parameter systems.

In Chapter 2, the basic elements of input design problem are brought together. Some important results, for example, Fisher information matrix, and properties of test signals are also stated.

In Chapter 3, the optimal input design problem for an autoregressive model with output power constraints is considered.

In Chapter 4, the problem of optimal input design for estimating part of the system parameters is studied. The equivalence theorem and a sequential design algorithm are extended for D_s-optimality design.

In Chapter 5, uniform input designs are considered and a comparison of uniform minimal designs with D-optimal designs is carried out.

In Chapter 6, optimal experimental design for parameter estimation of distributed parameter systems is discussed. Simple examples from parabolic and hyperbolic systems are used to illustrate the proposed design methodology.

Finally, Chapter 7 contains brief concluding remarks and suggestions for future research.

1.6 SUMMARY OF CONTRIBUTIONS

The major original contributions in this thesis are briefly summarized below:

In Chapter 3, the input design problem for estimating the parameters of an AR model with an output power constraint is studied. The optimal input frequencies can be obtained by solving a set of non-linear equations (Equations (3.4.1) and (3.4.2)) without recourse to optimization techniques involving the calculation of determinants (Main Result Section 3.4). It is illustrated in Section 3.5, that for low order models, at least, there is substantial reduction in computational complexity by using this method, as compared to the optimization of the information matrix. It is also observed that under mild physical constraints (Equation (3.4.2a)), the accuracy of parameter estimation for the open-loop design is the same as in the case of minimum variance control law. These results have been published by the author [N6].

The problem of optimal input design for D_s-optimality is studied in Chapter 4. First, the equivalence of D-optimal and G-optimal designs for dynamic systems (presented by Mehra [M4]) has been extended for the D_s-optimal case (Theorem 4.3.2). Using Theorem 4.3.2, a sequential design algorithm studied by Mehra [M4] and Zarrop [Z3] is extended for the D_s-optimality case (Section

4.5). Some examples are given to illustrate the algorithm (Section 4.6 and Appendix 4.1). It is also shown in the case of first order models, that D_s-optimal design is better than D-optimal design in the sense of achievable variance on the parameter estimate (Sections 4.4 and 4.6). These results have also been published by the author [Q2].

In Chapter 5 we look at the problem of comparing minimal uniform designs with a D-optimal design. In Section 5.3, we consider some examples to demonstrate the efficiency of minimal uniform designs. In Section 5.4, we carry out a detailed analysis of the efficiency of a two parameter system and illustrate the achievable lower bound of this efficiency.

A method to design optimal experiments for estimating the parameters of a general class of DPS is given in Section 6.2. The design procedure is illustrated by two specific examples. In Section 6.3 the optimal location problem of measurement sensors for estimating the velocity of propagation and the damping coefficient of a vibrating string are considered. The optimal measurements position given by Equation (6.3.1) corresponds to the antinodes of the string. In Section 6.4 we consider a heat diffusion process in which both the boundary perturbation and location of measurement sensors are treated as design variables. The joint optimal experimental conditons are given by Equations (6.4.26) and (6.4.27). These results have been published by the author [Q1].

CHAPTER 2

BASIC ELEMENTS
OF
INPUT DESIGN

CHAPTER 2: BASIC ELEMENTS OF INPUT DESIGN

2.1 INTRODUCTION

In this Chapter, we present some definitions and basic results of the input design problem which are relevant to the later part of the thesis. The main steps of the Fisher information matrix are derived, followed by its representation in frequency domain. Minimal properties of test signals and some commonly-used criteria of optimality are briefly discussed.

2.2 FISHER INFORMATION MATRIX

Consider a linear, time invariant, single-input single-output, discrete-time dynamic system described by:

$$y_k = \frac{B(z^{-1})}{A(z^{-1})} u_{k-d} + \frac{D(z^{-1})}{C(z^{-1})} e_k, \quad k = 1,2, \ldots, N \tag{2.2.1}$$

where:

$\{u_k\}$ and $\{y_k\}$ are the input and output sequences respectively,

d is the time delay from input to output, $\{e_k\}$ is a sequence of random variables, i.i.d. with variance σ^2, and:

$$A(z^{-1}) = 1 + a_1 z^{-1} + \ldots, + a_n z^{-n}$$

$$B(z^{-1}) = b_0 + b_1 z^{-1} + \ldots, + b_m z^{-m} \quad (b_0 \neq 0)$$

$$C(z^{-1}) = 1 + c_1 z^{-1} + \ldots, + c_q z^{-q}$$

$$D(z^{-1}) = d_0 + d_1 z^{-1} + \ldots, + d_r z^{-r} \quad (d_0 \neq 0)$$

It is assumed that all the polynomials, i.e. $A(z^{-1})$, $B(z^{-1})$, $C(z^{-1})$, $D(z^{-1})$ have roots inside the unit circle. It is also assumed that there are no pole-zero cancellations in Equation (2.2.1).

In this thesis, we also assume that the order of the system is known, i.e. n,m,q,r, and also d are known integers.

The parameter vector given by:

$$\beta^T = (a_1, \ldots, a_n, b_0, \ldots, b_m, c_1, \ldots, c_q, d_0, \ldots, d_r)$$

is to be estimated from input-output data.

The Fisher information matrix is described by the following expression:

$$M_\beta = E_{y/\beta}\left\{\left[\frac{\partial \log p(y/\beta)}{\partial \beta}\right]\left[\frac{\partial \log p(y/\beta)}{\partial \beta}\right]^T\right\} \quad (2.3.1)$$

where:

$p(y/\beta)$ denotes the likelihood function for the data,

$E_{y/\beta}$ denotes the conditional expectation over the data, given the parameters, and

y denotes the vector containing the N output values as components, i.e.

$$y \triangleq (y_1, y_2, \ldots, y_N)^T$$

If the noise sequence is assumed to be normally distributed, the log-likelihood function can be derived as: (Goodwin, Murdoch & Payne [G4])

$$L = \log p(y/\beta) = -\frac{1}{2} \sum_{k=1}^{N} \varepsilon_k^2 + \text{constant} \quad (2.3.2)$$

where:

N is the number of observations, and $\{\varepsilon_k\}$ is the residual sequenced defined by:

$$\varepsilon_k = \frac{C(z^{-1})}{D(z^{-1})} \{y_k - z^{-d} \frac{B(z^{-1})}{A(z^{-1})} u_k\} \tag{2.3.3}$$

and:

$$\frac{\partial L}{\partial \beta} = - \sum_{k=1}^{N} \varepsilon_k \frac{\partial \varepsilon_k}{\partial \beta} \tag{2.3.4}$$

Taking differentials of (2.3.3), we have:

$$\left.\begin{aligned} \frac{\partial \varepsilon_k}{\partial a_i} &= \frac{C(z^{-1})}{D(z^{-1})} \frac{B(z^{-1})}{A^2(z^{-1})} z^{-(d+i)} u_k, \quad i=1,2,\ldots,n \\ \frac{\partial \varepsilon_k}{\partial b_i} &= -\frac{C(z^{-1})}{D(z^{-1})} \frac{1}{A(z^{-1})} z^{-(d+i)} u_k, \quad i=0,\ldots,m \end{aligned}\right\} \tag{2.3.5}$$

$$\left.\begin{aligned} \frac{\partial \varepsilon_k}{\partial c_i} &= \frac{z^{-i}}{C(z^{-1})} \varepsilon_k, \quad i=1,\ldots,q \\ \frac{\partial \varepsilon_k}{\partial d_i} &= -\frac{z^{-i}}{D(z^{-1})} \varepsilon_k, \quad i=0,\ldots,r \end{aligned}\right\} \tag{2.3.6}$$

Note that:

(i) $\{\frac{\partial \varepsilon_k}{\partial a_i}\}$ and $\{\frac{\partial \varepsilon_k}{\partial b_i}\}$

do not depend on $\{\varepsilon_k\}$;

(ii) $\{\frac{\partial \varepsilon_k}{\partial c_i}\}$ and $\{\frac{\partial \varepsilon_k}{\partial d_i}\}$

are independent of $\{u_k\}$.

Substituting (2.3.4) into (2.3.1) and using (2.3.5) and (2.3.6), leads to the following simplified form of the information matrix:

$$M_\beta = \begin{bmatrix} M & 0 \\ 0 & R \end{bmatrix} \tag{2.3.7}$$

where the partition of M_β corresponds to a partition of β between the system parameters and the noise parameters, i.e.:

$$\beta^T = (\theta^T \,\vdots\, \theta_n^T)$$

where:

$$\theta^T = (a_1, \ldots, a_n, \; b_0, \ldots, b_m)$$

$$\theta_n^T = (c_1, \ldots c_q, \; d_0, \ldots, d_r)$$

For long data lengths ($N \to \infty$) the inverse of the information matrix represents the minimal achievable covariance for any unbiased estimator of the parameters (Cramer-Rao lower bound) (Nahi [N1]), i.e.:

$$\operatorname{cov} \hat{\beta} \geq M_\beta^{-1} = \begin{bmatrix} M^{-1} & 0 \\ 0 & R^{-1} \end{bmatrix} \tag{2.3.8}$$

and:

$$\left.\begin{aligned} \operatorname{cov} \hat{\theta} &\geq M^{-1} \\ \operatorname{cov} \hat{\theta}_n &\geq R^{-1} \end{aligned}\right\} \tag{2.3.9}$$

Equation (2.3.9) shows that the accuracy with which the noise parameters can be estimated is not influenced by the choice of input sequence.

Since the input sequence affects only the minimum variance bound of the system parameters, we shall consider the information matrix M for system parameters in the optimization procedure.

2.3 FREQUENCY DOMAIN REPRESENTATION OF INFORMATION MATRIX

For N large, it is more convenient to consider the asymptotic per sample information matrix defined by:

$$\bar{M} = \lim_{N\to\infty} \frac{1}{N} M$$

Following the derivation of Zarrop [Z3] or Goodwin and Payne [G6], the Equations (2.3.5) are written in the vector form:

$$\frac{\partial \varepsilon_k}{\partial \theta} = h(z^{-1}) \; u_k \tag{2.4.1}$$

where:

$$\left.\begin{aligned} h_i(z^{-1}) &= \frac{C(z^{-1})}{D(z^{-1})} \frac{B(z^{-1})}{A^2(z^{-1})} z^{-(d+i)}, \; i=1,2,\ldots,n \\ &= -\frac{C(z^{-1})}{D(z^{-1})} \frac{1}{A(z^{-1})} z^{-(d+i-n-1)}, \; i=n+1,\ldots,p \end{aligned}\right\} \tag{2.4.2}$$

$$\theta^T = (a_1,\ldots,a_n, b_0,\ldots,b_m)$$

$$p = n+m+1$$

Application of Parseval's theorem (Melsa & Sage [M6]) then gives:

$$\bar{M} = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \left(\frac{\partial \varepsilon_k}{\partial \theta}\right) \left(\frac{\partial \varepsilon_k}{\partial \theta}\right)^T$$

$$= \int_{-\pi}^{\pi} h(e^{j\omega})\, h^*(e^{j\omega})\, d\xi'(\omega) \qquad (2.4.3)$$

where:

$$h^T(e^{j\omega}) = [h_1(e^{j\omega}) \;\ldots\; h_p(e^{j\omega})]$$

$h^*(e^{j\omega})$ is the complex conjugate transpose of $h(e^{j\omega})$, and $\xi'(\omega)$ is the power spectral density of the input sequence $\{u_k\}$.

Note that: $\bar{M}$ does not depend on the time delay d.

In practice, it is more useful to work with the single-sided power distribution function $\xi(\omega)$ defined on $[0,\pi]$. Then Equation (2.4.3) can be written as:

$$\bar{M} = Re \int_{0}^{\pi} h(e^{j\omega})\, h^*(e^{j\omega})\, d\xi(\omega)$$

where:

$$\xi(\omega) = 2\xi'(\omega), \quad \omega \in (0,\pi)$$

$$= \xi'(\omega), \quad \omega = 0 \text{ or } \pi$$

In the experimental design literature, ξ is referred to as the design measure or the design.

<u>Definition</u>

Denote by Ξ the set of all ξ for which

$$\int_0^\pi d\xi(\omega) = 1.$$

This corresponds to normalized input power

If the input spectrum is discrete, then:

$$\bar{M} = Re \sum_{i=1}^{\ell} \lambda_i \, h(e^{j\omega_i}) \, h^*(e^{j\omega_i})$$

where:

λ_i corresponds to the proportion of power at frequency ω_i,

and:

$$\sum_{i=1}^{\ell} \lambda_i = 1$$

for normalized designs.

For properties of the information matrix, interested readers are referred to Mehra [M4] and Goodwin and Payne [G6].

2.4 PROPERTIES OF TEST SIGNALS

The minimal properties of test signals necessary for parameter identification have been studied by a number of authors, for example Aström and Bohlin [A2], Tse [T1], Aoki and Staley [A1]. A necessary condition for the p parameters to be identifiable is that the input spectrum consist of at least p/2 frequencies. Such an input is called persistently exciting by Aström and Bohlin [A2]. Rothenberg [R2] has shown that local parameter identifiability is equivalent to non-singularity of the information matrix.

The following theorem is due to Mehra [M4] and Goodwin and Payne [G6]:

Theorem 2.4.1

The information matrix $\bar{M}$ has the following properties:

(i) $\bar{M}$ is a real, symmetric, positive semi-definite matrix;

(ii) $\bar{M}$ is non-singular if the input consists of at least p/2 or (p+1)/2 frequencies, where p is the number of parameters;

(iii) $\bar{M}$ can be represented by a point in a p(p+1)/2 dimensional space;

(iv) The set of information matrices M corresponding to all normalized input designs is convex and closed.

(v) The set M is the convex hull of the set of information matrices corresponding to single frequency designs;

(vi) The optimal information matrix lies on the boundary of the convex hull.

∇∇∇

Payne [P2] exploited the special structure of the information matrix for SISO systems and further simplified the input design problem.

<u>Theorem 2.4.2</u> (Payne [P2])

For the SISO linear dynamic system represented by Equation (2.2.1), the information matrix $\bar{M}$ lies in a p dimensional space.

∇∇∇

This theorem implies a considerable reduction in the number of frequencies from $p(p+1)/2$ to, at most, p.

Zarrop [Z2] showed that, under certain conditions, the number of frequencies for the optimal design can be further reduced.

<u>Theorem 2.4.3</u> (Zarrop [Z2])

Consider the model given by Equation (2.2.1). If the system satisfies the following (hyperplane conditions):

(i) for input power constraint case

$$q = 0$$

$$m \geq n+r$$

(ii) for output power constraint case

$$q = 0$$

$$r = 0$$

If $A(z^{-1})$ and $C(z^{-1})$ have n_x roots in common, then the hyperplane conditions are:

(iii) for input power constraint case

$$q = n_x$$

$$m \geq n+r-n_x$$

(iv) for output power constaint case

$$q = n_x$$

$$r \leq n_x$$

Then the D-optimal design comprises p/2 frequencies (the minimum number of frequencies necessary to ensure identifiability).

∇∇∇

In general, the number of frequencies required for the existence of D-optimal designs is between p/2 and p.

2.5 CRITERIA OF OPTIMALITY

Some commonly used criteria of optimality employed for input design are:

(1) D-optimality

(2) A-optimality

(3) G-optimality

(1) D-Optimality:

This criterion involves choosing an input design to maximize det $\bar{M}$ or minimize det $\bar{M}^{-1}$. D-optimality is a parameter space criterion and it minimizes the generalized variance of the parameter estimates. An important advantage of D-optimality is that it is invariant under scale changes in the parameters and linear transformations of the output. If the distribution of $\hat{\theta}$ is assumed normal, D-optimality can be interpreted geometrically as the volume of the uncertainty ellipsoid, which we try to minimize with respect to the input. Figure 2.1 shows this criterion in two dimensions.

(2) A-Optimality:

This criterion minimizes the trace of $\bar{M}^{-1}$, i.e. choose an input to minimize the average variance of the parameter estimates. Certain authors, for example Goodwin et al. [G4], Zarrop [Z3] used tr $\bar{M}^{-1}W$, where W is a positive definite matrix.

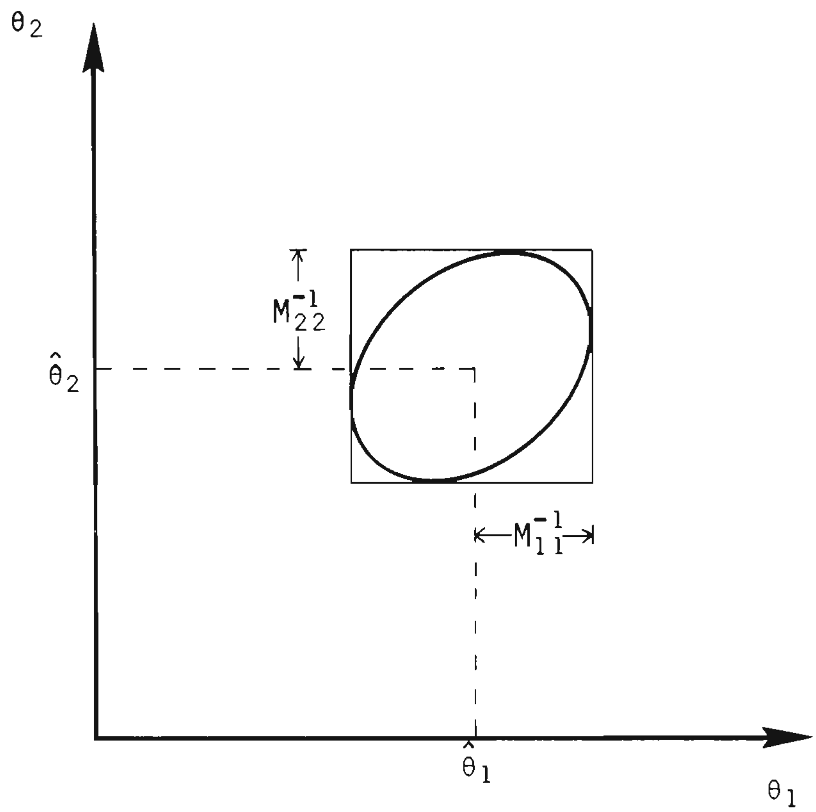


Fig. 2.1: Uncertainty ellipsoid in two dimensions

Some authors, for example Nahi and Wallis [N3], Aoki and Staley [A1], Nahi and Napjus [N2], Mehra [M1], used optimal control methods to maximize tr M. This criterion leads to a quadratic optimization problem which is easy to solve, but the inputs obtained are often unsuitable for parameter estimation. This has been pointed out by Goodwin [G2] and Zarrop and Goodwin [Z1].

(3) G-Optimality:

This is an output space criterion and is employed if the purpose of parameter identification is to predict accurately the output for different input signals. The input is then chosen such that the transfer function of the system is identified with the greatest accuracy. G-optimality is a minmax design — $\min_{\xi} \max_{\omega} d(\omega, \xi)$; i.e. choose an input design ξ to minimize the maximum in ω of $d(\omega, \xi)$, where $d(\omega, \xi)$ is the response dispersion (see Chapter 4 and Goodwin & Payne [G6]).

D-optimal and G-optimal designs have been shown to be equivalent by Kiefer and Wolfowitz [K3], and for dynamic systems by Mehra [M4].

For further details and discussion on the above criteria and other commonly used criteria, see Fedorov [F1] and Mehra [M4].

CHAPTER 3

OPTIMAL INPUT DESIGN FOR AUTOREGRESSIVE MODEL WITH OUTPUT POWER CONSTRAINT

CHAPTER 3: OPTIMAL INPUT DESIGN FOR AUTOREGRESSIVE MODEL WITH OUTPUT POWER CONSTRAINT

3.1 INTRODUCTION

In this Chapter, the output power constraint problem of optimal input design for an autoregressive model is considered. It is shown that, besides the minimum variance control law design, an open-loop D-optimal design exists consisting of sinusoidal input test signals. The optimal input frequencies can be obtained by solving a set of non-linear equations without recourse to optimization techniques involving the calculation of determinants (Sec.3.4). Furthermore, it is shown that, under mild physical constraints, the accuracy of parameter estimation for the open-loop design is the same as in the case of minimum variance control law design. Several examples are given to illustrate the design methodology (Sec.3.5).

3.2 EXISTENCE OF OPTIMAL INPUT DESIGNS

In this Section, we show the existence of a solution to the problem of input design for parameter estimation.

First, we state Caratheodory's Theorem:

Theorem 3.2.1: (Fedorov [F1])

Each point s^0 in the convex hull S^0 of any subset S, of the n dimensional space can be represented in the form:

$$s^0 = \sum_{i=1}^{n+1} \alpha_i \, s_i$$

where:

$$\alpha_i \geq 0 \;,\; \sum_{i=1}^{n+1} \alpha_i = 1 \;,\; s_i \in S$$

If s^0 is a boundary point of the set S^0, then α_{n+1} can be set equal to zero.

$\nabla\nabla\nabla$

From Caratheodory's Theorem and property (iii) of Theorem (2.4.1) it follows that $\overline{M}$ can be represented as:

$$\overline{M} = \sum_{i=1}^{k} \lambda_i \, \overline{M}(\omega_i)$$

where:

$$k \leq \frac{p(p+1)}{2} + 1$$

$$0 \leq \lambda_i \leq 1$$

$$\sum_{i=1}^{k} \lambda_i = 1$$

This implies the existence of input designs comprising a finite number of frequencies (Mehra [M4]).

Following Section 2.4, D-optimal input designs exist for frequencies lying between p/2 and p.

Now we show the existence of D-optimal input designs for an autoregressive model.

Theorem 3.2.2:

For an autoregressive model, represented by Equation (3.3.1) or by:

$$y_k = \frac{b}{A(z^{-1})} u_{k-d} + \frac{1}{A(z^{-1})} e_k \tag{3.2.1}$$

with a constraint imposed on the output power.

There exists a D-optimal input design comprising p/2 frequencies.

Proof:

The autoregressive model satisfies the hyperplane condition (iv) of Theorem 2.4.3.

Hence the result follows.

$\nabla\nabla\nabla$

In the next Section, we formulate the input design problem, and present a method to obtain the optimal input frequencies.

3.3 PROBLEM FORMULATION

Consider the following AR model:

$$y_k = a_1 y_{k-1} + \dots + a_n y_{k-n} + bu_{k-1} + e_k \tag{3.3.1}$$

where $\{e_k\}$ is a white noise sequence having a Gaussian distribution with variance σ^2. We consider the case where the variance of the system output $\{y_k\}$, cannot be greater than some prescribed value (say, W), i.e.:

$$W \geq E\{y_k^2\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} dFy(\omega) \tag{3.3.2}$$

where $Fy(\omega)$ is the spectral distribution of the output sequence $\{y_k\}$.

The constraint is chosen by physical conditions. In the case where feedback is present, the minimum value of $E\{y_k^2\}$ can be reduced to σ^2 (Ng et al. [N4]).

Following Ng et al., [N4], the average information matrix for the system parameters $(a_1,\dots,a_n,b)$ is given by:

$$\bar{M} = E_{y/\theta}\{\frac{1}{N\sigma^2} \sum_{k=1}^{N} [\frac{\partial \varepsilon_k}{\partial \theta}]^T [\frac{\partial \varepsilon_k}{\partial \theta}]\} \tag{3.3.3}$$

where:

$$\theta = (a_1, \dots, a_n, b)^T$$

$$= (\alpha^T, b)^T$$

$E_{y/\theta}$ denotes the conditional expectation over the data given the parameters and ε_k is the residual sequence given by:

$$\varepsilon_k = y_k - a_1 y_{k-1}, \ldots, - a_n y_{k-n} - bu_{k-1} \tag{3.3.4}$$

Thus:

$$\frac{\partial \varepsilon_k}{\partial a_i} = -y_{k-i} , \quad i=1,\ldots, n \tag{3.3.5}$$

$$\frac{\partial \varepsilon_k}{\partial b} = -u_{k-1} \tag{3.3.6}$$

Substitute (3.3.5), (3.3.6) into (3.3.3), which yields:

$$\bar{M} = E\frac{1}{\sigma^2}\left[\begin{array}{ccc|c} y^2_{k-1} & \cdots & y_{k-1}\, y_{k-n} & y_{k-1}\, u_{k-1} \\ \vdots & & \vdots & \vdots \\ y_{k-n}\, y_{k-1} & \cdots & y^2_{k-n} & y_{k-n}\, u_{k-1} \\ \hline u_{k-1}\, y_{k-1} & \cdots & u_{k-1}\, y_{k-n} & u^2_{k-1} \end{array}\right] \tag{3.3.7}$$

$$= \frac{1}{\sigma^2} \left[\begin{array}{c|c} F & G \\ \hline G^T & H \end{array} \right] \qquad (3.3.8)$$

where the partitions correspond to the partitioning of θ into $(\alpha^T , b)^T$

We now define $\{\rho_i\}$ as follows:

$$E \{y_{k-j} \; y_{k-\ell}\} \triangleq \rho_{|j-\ell|} \qquad (3.3.9)$$

Substitute (3.3.9) into (3.3.8) and using (3.3.1), we have:

$$F = \begin{bmatrix} \rho_0 & \rho_1 & \cdot & \cdot & \cdot & \rho_{n-1} \\ \rho_1 & \ddots & & & & \cdot \\ \cdot & & \ddots & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & \ddots & \rho_1 \\ \rho_{n-1} & \cdot & \cdot & \cdot & \rho_1 & \rho_0 \end{bmatrix} \qquad (3.3.10)$$

$$G = \frac{1}{b} (V - F\alpha) \qquad (3.3.11)$$

$$H = \frac{1}{b^2} (\rho_0 - 2\alpha^T V + \alpha^T F\alpha - \sigma^2) \qquad (3.3.12)$$

where V is given by:

$$V^T = [\rho_1, \rho_2, \ldots, \rho_n] \qquad (3.3.13)$$

From (3.3.8) we have:

$$\log \det \bar{M} = \log \det(F) + \log(H - G^T F^{-1} G) - (n+1) \log (\sigma^2) \tag{3.3.14}$$

Substitution of (3.3.11) to (3.3.13) into (3.3.14) yields:

$$\log \det \bar{M} = \log \det(F) + \log(\rho_0 - \sigma^2 - V^T F^{-1} V) - \log(b^2) - (n+1) \log(\sigma^2) \tag{3.3.15}$$

We now state the following theorem:

<u>Theorem 3.3.1</u>: (Ng et al. [N4])

For the system described by Equation (3.3.1), and satisfying:

(i) the feed back law is causal;

(ii) a constraint described by Equation (3.3.2) is imposed;

(iii) the determinant of the information matrix is used as a design criterion;

then an optimal design exists comprising a minimum variance control law, together with a white test signal.

Proof:

log det $\bar{M}$ of Equation (3.3.15) is mamixized if and only if:

$$\rho_1 = \rho_2 = \ldots = \rho_n = 0$$

Note that this condition can be achieved when $\{y_k\}$ is an uncorrelated sequence and this is true if the input $\{u_k\}$ is chosen to satisfy:

$$u_k = \frac{1}{b}(a_1 y_{k-1} + a_2 y_{k-2} + \ldots + a_n y_{k-n}) + s_k \tag{3.3.16}$$

where:

s_k is a white external test sequence.

Equation (3.3.16) is the minimum variance control law for the model of Equation (3.3.1).

∇∇∇

In the next Section we present a method to obtain a minimal D-optimal input design without using optimization techniques involving the calculation of determinants.

3.4 A METHOD TO OBTAIN OPTIMAL INPUT FREQUENCIES

Theorem 3.4.1:

For the autoregressive model described by Equation (3.3.1) and satisfying:

(i) an output power constraint described by (3.3.2) is imposed;

(ii) the determinant of the information matrix is used as a design criterion;

then there is a set of constrained non-linear equations whose solution is the D-optimal input frequencies.

Proof:

With input parameterized as:

$$u_k = \sum_{i=1}^{\ell} m_i' \cos \omega_i k$$

where:

$$m_i' \geq 0 \ , \ i=1 \ , \ldots \ \ell;$$

$$\ell = (n+1)/2 \text{ or } (n+2)/2$$

and the frequencies $\omega_i \in [0,2\pi]$ are distinct, and are chosen to satisfy:

$$\rho_0 = W$$

Using (3.3.1), (3.3.7), (3.3.9), and applying the condition:

$$\rho_1 = \rho_2 = \ldots \ \rho_n = 0,$$

we arrive at the following equations for optimality:

$$\rho_0 = \sum_{i=1}^{\ell} \frac{m_i}{f(\omega_i)} + C_0 = W \tag{3.4.1}$$

$$\left.\begin{array}{l}
\rho_1 = \displaystyle\sum_{i=1}^{\ell} \frac{m_i \cos\omega_i}{f(\omega_i)} + C_1 = 0 \\
\rho_2 = \displaystyle\sum_{i=1}^{\ell} \frac{m_i \cos 2\omega_i}{f(\omega_i)} + C_2 = 0 \\
\cdot \\
\cdot \\
\cdot \\
\rho_n = \displaystyle\sum_{i=1}^{\ell} \frac{m_i \cos n\omega_i}{f(\omega_i)} + C_n = 0
\end{array}\right\} \tag{3.4.2a}$$

where:

$$m_i' = m_i'^2 \, b^2$$

$$f(\omega_i) = A(e^{j\omega_i}) \,.\, A(e^{-j\omega_i}) ; \quad i=1, \ldots, \ell$$

$$C_k = \frac{\sigma^2}{2\pi} \int_{-\pi}^{\pi} \frac{e^{-jk\omega}}{A(e^{j\omega})A(e^{-j\omega})} d\omega; \quad k=0, 1, \ldots, n$$

$$C_0 \geq C_j; \quad j=1, 2, \ldots, n$$

$$A(z) = 1 - a_1 z - a_2 z^2 \ldots, -a_n z_n$$

Physical consideration (cf. Söderström [S6]) also requires:

$$E \left\{\frac{bu_{k-i}}{A(z)}\right\} \geq E \left\{\frac{bu_{k-i}}{A(z)} \; \frac{bu_{k-j}}{A(z)}\right\}$$

which is equivalent to:

$$W - C_0 \geq |C_j| , \; j=1, \ldots, n. \tag{3.4.2b}$$

∇∇∇

A solution to (3.4.1) and (3.4.2) could be obtained by numerical techniques or, alternatively, by writing (3.4.1) and (3.4.2a) in matrix form, and, checking that conditions (3.4.2b) are satisfied, we have:

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \cos\omega_1 & \cos\omega_2 & & \cos\omega_\ell \\ \cos 2\omega_1 & & & \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ \cos n\omega_1 & \cdots & & \cos n\omega_\ell \end{bmatrix} \begin{bmatrix} \dfrac{m_1}{f(\omega_1)} \\ \cdot \\ \cdot \\ \dfrac{m_\ell}{f(\omega_\ell)} \end{bmatrix} = \begin{bmatrix} W - C_0 \\ - C_1 \\ \cdot \\ \cdot \\ \cdot \\ - C_n \end{bmatrix} \tag{3.4.3}$$

Since the rows of the matrix on the lefthand side of (3.4.3) are independent, we could first solve for $\frac{m_i}{f(\omega_i)}$, $i=1,\ldots,\ell$ in terms of ω_i, $i=1,\ldots,\ell$ using the first ℓ equations. Then substitute into the remaining $n+1-\ell$ equations ($n+1-\ell \leq \ell$) to obtain a solution (or solutions) for ω_i, $i=1,\ldots,\ell$. Back substitute ω_i, $i=1,\ldots,\ell$ into the first ℓ equations, we could solve for m_i, $i=1,\ldots,\ell$. Finally, the conditions $m_i \geq 0$, $i=1,\ldots,\ell$ are checked.

Note that:

(i) using either the minimum variance control law or applying (3.4.1) and (3.4.2), the same maximum value of det $\overline{M}$ is achieved, i.e.:

$$\det \overline{M} = \frac{\rho_0{}^n(\rho_0 - \sigma^2)}{b^2 \sigma^{2(n+1)}}$$ which is also independent of the parameters a_i, $i=1,\ldots n$;

(ii) there is a difference in regard to the minimum value of the constraint W.

For the minimum variance control case, the minimum W is σ^2.

In the open-loop input design, the minimum W is $C_0 + \mathrm{Max}\,|C_i|$, $i=1,\ldots n$ which is dependent of the parameters a_i, $i=1,\ldots n$.

If W is chosen greater than or equal to $2C_0$, then constraint (3.4.2b) will always be satisfied.

The complex integral C_k, $k=0,1,\ldots,n$ is evaluated as shown in Appendix 3.1. Numerical methods to evaluate C_k can be found in Jury [J1] and Aström et al. [A6].

3.5 EXAMPLES

Example 1:

Consider the first order example:

$$y_k = a_1 y_{k-1} + bu_{k-1} + e_k$$

where:

$$-1 < a < 1, \quad E\, e_k^2 = \sigma^2 \text{ and } E\, y_k^2 = W \qquad (3.5.1)$$

Detailed analysis of this example has been given in Söderström [S6]).

Note that optimal design requires at most one input frequency; following (3.4.1) and (3.4.2a), we arrive at the following equations:

$$\rho_0 = \frac{m_1}{1 + a^2 - 2a\cos\omega} + \frac{\sigma^2}{1 - a^2} = W \qquad (3.5.2)$$

$$\rho_1 = \frac{m_1 \cos\omega}{1 + a^2 - 2a\cos\omega} + \frac{a\,\sigma^2}{1 - a^2} = 0 \qquad (3.5.3)$$

Substituting $\dfrac{m_1}{1 - a^2 - 2a\cos\omega}$ from (3.5.2) into (3.5.3) yields:

$$\cos\omega^{\dagger} = \frac{-a\,\sigma^2}{W(1-a^2)-\sigma^2} \qquad (3.5.4)$$

and

$$\det \overline{M} = \frac{W(W-\sigma^2)}{b^2\sigma^4} \qquad (3.5.5)$$

Alternatively, following Section 3.3 and using (3.5.1), we arrive at the following information matrix for the above system:

$$\bar{M} = \frac{1}{\sigma^2} \begin{bmatrix} W & Q_1 \\ Q_1 & Q_2 \end{bmatrix} \qquad (3.5.6)$$

where:

$$Q_1 = -\frac{\cos\omega - a}{b} \cdot \left(W - \frac{\sigma^2}{1-a^2}\right)$$

$$Q_2 = \frac{1 + a^2 - 2a\cos\omega}{b^2} \cdot \left(W - \frac{\sigma^2}{1-a^2}\right)$$

$$\det \bar{M} = \frac{1}{b^2\sigma^4} \cdot \left(W - \frac{\sigma^2}{1-a^2}\right) \cdot \left[W(1+a^2-2a\cos\omega) - (\cos\omega - a)^2\left(W - \frac{\sigma^2}{1-a^2}\right)\right] \qquad (3.5.7)$$

To maximize det $\bar{M}$, we differentiate (3.5.7) w.r.t. ω, i.e.:

$$\frac{d \det \bar{M}}{d\omega} = 0 \Rightarrow \cos\omega^{\dagger} = \frac{-a\,\sigma^2}{W(1-a^2) - \sigma^2}$$

back substitution into (3.5.7) gives:

$$\det \bar{M}^{\dagger} = \frac{W(W-\sigma^2)}{b^2\sigma^4}$$

Physical constraint (3.4.2b) requires $W \geq \frac{\sigma^2}{1 - |a|}$ which is equivalent to constraint $|\cos\omega^{\dagger}| \leq 1$.

From this example, we observe that, using the present approach, we save a considerable amount of computational effort in the search for optimal input frequencies.

Example 2:

Consider the following second order example:

$$y_k = a_1 y_{k-1} + a_2 y_{k-2} + bu_{k-1} + e_k \qquad (3.5.8)$$

with values of a_1, a_2 such that $A(z^{-1})$ have roots inside the unit circle, $E\, e_k^2 = \sigma^2$, and $E\, y_k^2 = W$.

Following (3.4.1) and (3.4.2a), we have:

$$\rho_0 = \frac{m_1}{f(\omega_1)} + \frac{m_2}{f(\omega_2)} + C_0 = W \qquad (3.5.9)$$

$$\rho_1 = \frac{m_1 \cos\omega_1}{f(\omega_1)} + \frac{m_2 \cos\omega_2}{f(\omega_2)} + C_1 = 0 \qquad (3.5.10)$$

$$\rho_2 = \frac{m_1 \cos 2\omega_1}{f(\omega_1)} + \frac{m_2 \cos 2\omega_2}{f(\omega_2)} + C_2 = 0 \qquad (3.5.11)$$

where:

$$f(\omega) = (1+a_1^2+a_2^2) - 2a_1(1-a_2)\cos\omega - 2a_2\cos 2\omega$$

Physical constraints require:

$$W - C_0 \geq |C_j| \ , \ j=1,2 \qquad (3.5.12)$$

C_k, k=0,1,2 is evaluated using the method in Appendix 3.1.

$$C_0 = \frac{1-a_2}{1+a_2} \ \frac{\sigma^2}{1-(a_1{}^2+2a_2) + a_2{}^2}$$

$$C_1 = \frac{a_1}{1+a_2} \ \frac{\sigma^2}{1-(a_1{}^2+2a_2) + a_2{}^2}$$

$$C_2 = \frac{a_1{}^2 + a_2\ (1-a_2)}{1 + a_2} \ \frac{\sigma^2}{1-(a_1{}^2+2a_2) + a_2{}^2}$$

setting $a_1 = 0.5$, $a_2 = -0.1$, $b = 0.5$, we obtain $C_0 = 1.273$, $C_1 = 0.578$, $C_2 = 0.16$; choosing $W \geq 2C_0$ so that the physical constraints are satisfied, say W=5, and arbitrarily choosing $\omega_1 = \pi/4$, we arrive at the following D-optimal values:

$$\omega_1 = \pi/4; \qquad m_1 = 0.6707 \text{ or } m_1' = 1.6379$$

$$\omega_2 = 2.3026; \qquad m_2 = 4.7111 \text{ or } m_2' = 4.3410$$

We remark that, for this particular example, there exists an infinite number of open-loop D-optimal designs. This is clearly shown by substituting $\frac{m_1}{f(\omega_1)}$, $\frac{m_2}{f(\omega_2)}$ from (3.5.9) and (3.5.10) into (3.5.11), which then yields an

equation with two unknowns. Hence, by arbitrarily choosing ω_1, we obtained the corresponding ω_2. We also observe that the choice of W will influence the D-optimal frequencies.

3.6 CONCLUSION

A method to obtain an optimal input design for an autoregressive model with output power constraint has been presented. Optimal input frequencies have been obtained by solving a set of non-linear equations without recourse to optimization techniques involving the calculation of determinants. It has been illustrated that, for low order models, the reduction in computational complexity, by using this method, as compared to the optimization of the information matrix, is substantial. It has also been observed that open-loop design is more restrictive than closed-loop design as physical constraints have to be satisfied.

APPENDIX 3.1: EVALUATION OF INTEGRAL

$$C_k = \frac{\sigma^2}{2\pi} \int_{-\pi}^{\pi} \frac{e^{-jk\omega}}{A(e^{j\omega})\; A^*(e^{j\omega})}\, d\omega$$

Our goal is to apply the residue theorem to evaluate the integral:

$$C_k = \frac{\sigma^2}{2\pi} \int_{-\pi}^{\pi} \frac{e^{-jk\omega}}{A(e^{j\omega})\; A^*(e^{j\omega})}\, d\omega \tag{A3.1.1}$$

(A3.1.1) can be identified as the parametrized form of a contour integral, $\int_C F(z)\, dz$, of some complex function F along the unit circle $C : |z| = 1$ (see, for example, Saff & Snider [S9]). To establish this identification we parametrize C by:

$$z = e^{j\omega} \tag{A3.1.2}$$

$$\therefore \qquad \frac{dz}{d\omega} = j\, e^{j\omega} = jz$$

so that:

$$d\omega = \frac{1}{j}\, \frac{dz}{z} \tag{A3.1.3}$$

Making the substitutions (A3.1.2) and (A3.1.3) in (A3.1.1), we get:

$$C_k = \frac{\sigma^2}{2\pi j} \int_C F(z)\, dz$$

where the new integral F is

$$F(z) = \frac{z^{-k}}{A(z)\; A(z^{-1})}\, \frac{dz}{z}$$

Rewrite:

$$C_k = \frac{\sigma^2}{2\pi j}\, I_k \tag{A3.1.4}$$

where:

$$I_k = \int_C \frac{z^{-k}}{z\, A(z)\, A(z^{-1})}\, dz$$

Next, we apply the residue theorem, and find the residues of F at its poles which lie inside the unit circle; this yields:

$$I_k = 2\pi j \sum_i \text{residues} \qquad \text{(A3.1.5)}$$

Substituting (A3.1.5) in (A3.1.4) gives:

$$C_k = \frac{\sigma^2}{2\pi j} \{2\pi j \sum_i \text{res}\}$$

$$= \sigma^2 \sum_i \text{res}$$

∇∇∇

Using the procedure mentioned above, we show the calculations for evaluation of C_k, k=0,1, ..., n.

Consider Example 1 (Section 3.5), where:

$$C_k = \frac{\sigma^2}{2\pi j} \int_C \frac{z^{-k}}{zA(z)\, A(z^{-1})}\, dz \quad k=0,1$$

Since: $A(z^{-1}) = 1 - a_1 z^{-1}$

then: $z\, A(z^{-1}) = z - a_1$

and: $A(z) = 1 - a_1 z$

$$= -a_1 (z - \frac{1}{a_1})$$

Therefore the denominator inside the integral can be written as:

$$z\,A(z^{-1})\,A(z) \;=\; -a_1(z-a_1)(z-\frac{1}{a_1})$$

For k=0,

$$C_0 \;=\; \frac{\sigma^2}{2\pi j}\,I_0$$

where:

$$I_0 \;=\; \oint_C \frac{1}{-a_1(z-a_1)(z-\frac{1}{a_1})}\,dz$$

$$\text{residue}\,(z = a_1) \;=\; \frac{1}{1-a_1{}^2}$$

$$\therefore \qquad I_0 \;=\; 2\pi j\,\{\frac{1}{1-a_1{}^2}\}$$

and: $$C_0 \;=\; \frac{\sigma^2}{1-a_1{}^2}$$

For k=1,

$$C_1 \;=\; \frac{\sigma^2}{2\pi j}\,I_1$$

where:

$$I_1 \;=\; \oint_C \frac{z}{-a_1(z-a_1)(z-\frac{1}{a_1})}\,dz$$

$$\text{residue } (z = a_1) = \frac{a_1}{1-a_1{}^2}$$

$$\therefore \qquad C_k = \frac{a_1 \sigma^2}{1-a_1{}^2}$$

Now, consider Example 2 (Section 3.4), where:

$$C_k = \frac{\sigma^2}{2\pi j} \int_C \frac{z^{-k}}{zA(z)\ A(z^{-1})} dz; \quad k=0,1,2$$

Since: $A(z^{-1}) = 1 - a_1 z^{-1} - a_2 z^{-2}$

then: $z^2\ A(z^{-1}) = z^2 - a_1 z - a_2$

Factorizing gives:

$$z^2\ A(z^{-1}) = (z-P_1)(z-P_2)$$

where:

$$a_1 = P_1 + P_2$$

$$a_2 = -P_1 P_2$$

Now:

$$A(z) = 1 - a_1 z - a_2 z^2$$

$$= -a_2 (z^2 + \frac{a_1}{a_2} z - \frac{1}{a_2})$$

and factorizing gives:

$$A(z) = -a_1(z-Q_1)(z-Q_2)$$

It is easy to show that:

$$Q_1 = \frac{1}{P_1}$$

$$Q_2 = \frac{1}{P_2}$$

Finally, writing $z^2\ A(z^{-1})\ A(z)$ as:

$$z^2\ A(z^{-1})\ A(z) = -a_2(z-P_1)(z-P_2)(z-\frac{1}{P_1})(z-\frac{1}{P_2})$$

For k=0,

$$C_0 = \frac{\sigma^2}{2\pi j} I_0$$

where:

$$I_0 = \int_C \frac{z}{-a_2(z-P_1)(z-P_2)(z-\frac{1}{P_1})(z-\frac{1}{P_2})} dz$$

$$\text{residue } (z = P_1) = -\frac{P_1{}^2 P_2}{a_2(P_1-P_2)(1-P_1{}^2)(1-P_1P_2)}$$

$$\text{residue } (z = P_2) = \frac{P_1 P_2{}^2}{a_2(P_1-P_2)(1-P_1P_2)(1-P_2{}^2)}$$

$$\Sigma \text{ residues} = \frac{1 + P_1P_2}{1 - P_1P_2} \quad \frac{1}{1 - (P_1{}^2+P_2{}^2) + P_1{}^2P_2{}^2}$$

Using $a_1 = P_1 + P_2$, $a_2 = -P_1P_2$, and substituting P_1 and P_2 in terms of a_1 and a_2 in the above expression, yields:

$$\Sigma \text{ residues} = \frac{1 - a_2}{1 + a_2} \quad \frac{1}{1 - (a_1{}^2+2a_2) + a_2{}^2}$$

Finally, we obtain:

$$C_0 = \frac{1 - a_2}{1 + a_2} \quad \frac{\sigma^2}{1 - (a_1{}^2+2a_2) + a_2{}^2}$$

For k=1,

$$C_1 = \frac{\sigma^2}{2\pi j} I_1$$

where:

$$I_1 = \int_C \frac{z^2}{-a_2(z-P_1)(z-P_2)(z-\frac{1}{P_1})(z-\frac{1}{P_2})} dz$$

Taking residues at P_1, P_2:

$$\text{residue } (z = P_1) = -\frac{P_1{}^3P_2}{a_2(P_1-P_2)(1-P_1{}^2)(1-P_1P_2)}$$

$$\text{residue } (z = P_2) = \frac{P_1P_2{}^3}{a_2(P_1-P_2)(1-P_2{}^2)(1-P_1P_2)}$$

$$\therefore \qquad \Sigma \text{ residues} = \frac{P_1 + P_2}{1 - P_1 P_2} \; \frac{1}{1 - (P_1^2 + P_2^2) + P_1^2 P_2^2}$$

Substituting for P_1, P_2 using $a_1 = P_1 + P_2$, $a_2 = -P_1 P_2$, yields:

$$\Sigma \text{ residues} = \frac{a_1}{1 + a_2} \; \frac{1}{1 - (a_1^2 + 2a_2) + a_2^2}$$

Finally, we obtain:

$$C_1 = \frac{a_1}{1 + a_2} \; \frac{\sigma^2}{1 - (a_1^2 + 2a_2) + a_2^2}$$

For k=2, the integral C_2 is evaluated similarly.

▽▽▽

CHAPTER 4

OPTIMAL INPUT DESIGN FOR PARAMETER ESTIMATION - THE D_s-OPTIMALITY CASE

CHAPTER 4: OPTIMAL INPUT DESIGN FOR PARAMETER ESTIMATION - THE D_s-OPTIMALITY CASE

4.1 INTRODUCTION

In this Chapter, the optimal input design problem for estimating s out of p parameters is considered.

It is evident that in designing identification experiments for different purposes, different design criteria should be used. For example, if the design is to minimize the covariance of the parameter estimator, a scalar function of the parameter covariance matrix (or the inverse of the information matrix M^{-1}) is a good choice. On the other hand, if the purpose of an identification experiment is to predict accurately the output sequence, then it is more reasonable to cost the output variance. A number of input design criteria have been suggested in Mehra [M4]. For our purpose, we consider only two of them:

(a) A design $\xi^{\dagger}$ is called D-optimal if it maximizes det $\bar{M}(\xi)$; and

(b) A design $\xi^{\dagger}$ is called G-optimal if it minimizes $\max_{\omega} d(\omega, \xi^{\dagger})$.

For a definition of $d(\omega, \xi)$, see Equation (4.2.3) and Goodwin [G6].

Excluding trivial cases, it is obvious that optimal design which satisfies all criteria does not exist. However, the two criteria D_s-optimality and G_s-optimality, which compare the results of an experiment in two different spaces, are strictly related to each other (see Atwood [A5], Fedorov [F1], Kiefer [K3], Mehra [M4] and Zarrop [Z3]). The equivalence of these two design criteria has proved to be extremely important. It is fundamental to most D-optimal design algorithms and many computational improvements in both linear regression problems and dynamic system design problems.

In Section 2, some definitions and terminology are established. We extend the proof of the equivalence for D_s-optimality and G_s-optimality in Section 3. The convergence of a sequential design algorithm is proven and illustrated by simple examples (Section 4.5; Appendix 4.1).

4.2 PROBLEM STATEMENT

Consider a linear, time invariant, SISO, discrete time dynamic system described by Equation (2.2.1).

The average information matrix for system parameters $a_1, \ldots, a_n$ and $b_0, \ldots, b_m$ is given by (see Chapter 2):

$$\bar{M}(\xi) = Re \int_0^{\pi} h(e^{j\omega})\ h^*(e^{j\omega})\ d\xi(\omega) \qquad (4.2.1)$$

where:

$$p = m + n + 1$$

and

ξ is a normalized design, i.e. $\int_0^{\pi} d\xi(\omega) = 1$

We assume that:

$\bar{M}(\xi)$ is non-singular, i.e. the design ξ is persistently exciting.

Since we are interested in estimating part of the system parameters, we partition the information matrix and its inverse into:

$$\bar{M}(\xi) = \left[\begin{array}{c|c} M_1(\xi) & M_2(\xi) \\ \hline M_2^T(\xi) & M_3(\xi) \end{array}\right]$$

and

$$\overline{M}^{-1}(\xi) = \left[\begin{array}{c|c} M^{(1)}(\xi) & M^{(2)}(\xi) \\ \hline M^{(2)T}(\xi) & M^{(3)}(\xi) \end{array}\right]$$

where:

$M_1(\xi)$ and $M^{(1)}(\xi)$ are s x s matrices, and

$M_3(\xi)$ and $M^{(3)}(\xi)$ are (p-s) x (p-s) matrices.

Now define:

(1) $M_s(\xi) = [M^{(1)}(\xi)]^{-1}$

$$= M_1(\xi) - M_2(\xi)\, M_3^{-1}(\xi)\, M_2^T(\xi) \tag{4.2.2}$$

and it follows:

$$\det M_s(\xi) = \det[M_1(\xi) - M_2(\xi)\, M_3^{-1}(\xi)\, M_2^T(\xi)]$$

$$= \frac{\det \overline{M}(\xi)}{\det M_3(\xi)} \tag{4.2.2a}$$

(2) $\xi^\dagger$ yields a global D_s-optimal design if:

$$\det M_s(\xi^\dagger) = \max_{\xi} \det M_s(\xi).$$

(3) $\xi^\dagger$ yields a local D-optimal design if:

$$\frac{\partial}{\partial\alpha} \log \det M_s[(1-\alpha)\xi^\dagger + \alpha\xi]\Big|_{\alpha=0} \leq 0 \text{ for all } \xi.$$

(4) the generalised variance $d(\omega,\xi)$, which in physical terms, can be interpreted as the ratio of the variance of the system frequency response to the noise power at frequency ω (see Goodwin and Payne [G6]) in the case $s \leq p$ to be:

$$d_s(\omega,\xi) = h^*(e^{j\omega})\ \bar{M}^{-1}(\xi)\ h(e^{j\omega})$$
$$- h^{(2)*}(e^{j\omega})\ M_3^{-1}(\xi)\ h^{(2)}(e^{j\omega}) \qquad (4.2.3)$$

where:

$$h^T(e^{j\omega}) = [h_1(e^{j\omega}), \ \ldots, \ h_s(e^{j\omega}) \ \vdots \ h_{s+1}(e^{j\omega}), \ \ldots, \ h_p(e^{j\omega})]$$
$$= [h^{(1)T}(e^{j\omega}) \ \vdots \ h^{(2)T}(e^{j\omega})]$$

In the next Section we prove the equivalence theorem for the s out of p parameter case.

4.3 EQUIVALENCE THEOREM

Before we proceed to state the main result, we first show that the local minimum of optimizing det $M_s(\xi)$ is also the global minimum.

Theorem 4.3.1:

Consider $\xi^o \in \Xi$

and let $\xi = (1-\alpha)\xi^\dagger + \alpha\xi^o \in \Xi$, $\alpha \in [0,1]$

Then:

(1) $\xi^\dagger$ maximizes det $M_s(\xi)$ for all $\xi \in \Xi$

(2) $\frac{\partial}{\partial\alpha}$ log det $M_s[(1-\alpha)\xi^\dagger + \alpha\xi^o] \leq 0$

are equivalent.

Proof:

Clearly, (1) $\rightarrow$ (2).

To show that:

(2) $\rightarrow$ (1),

we first state two matrix lemmas without proof.

Lemma 4.3.1 (Kiefer [K1])

For $i = 1, 2, \ldots, r$,

let C_i be $s \times (p-s)$,

and let D_i be positive definite symmetric

$s \times s$ matrices,

and suppose $\alpha_i > 0$, $\sum_{i=1}^{r} \alpha_i = 1$,

then:

$$\left[\sum_{i=1}^{r} \alpha_i C_i\right] \left[\sum_{i=1}^{r} \alpha_i D_i\right]^{-1} \left[\sum_{i=1}^{r} \alpha_i C_i^T\right] \leq \sum_{i=1}^{r} \alpha_i C_i D_i^{-1} C_i^T$$

with equality, if and only if the matrix C_i, D_i^{-1} is the same for all i.

Lemma 4.3.2 (Fedorov [F1])

If A and B are non-negative definite symmetric s x s matrices, then:

log det $(\alpha A + (1-\alpha)B)$ is concave for $\alpha \in [0,1]$ and is strictly concave unles $A = B$ or $A + B$ is singular.

Now assume that $\xi^\dagger$ does not maximize det $M_s(\xi)$ for all $\xi \in \Xi$, then there exist $\xi^o \in \Xi$ such that:

$$\log \det M_s(\xi^o) - \log \det M_s(\xi^\dagger) \geq 0$$

Let $\xi = (1-\alpha)\xi^\dagger + \alpha\xi^o \in \Xi$, then:

$$\begin{aligned} M_s(\xi) &= [M^{(1)}(\xi)]^{-1} \\ &= M_1(\xi) - M_2(\xi)\, M_3^{-1}(\xi)\, M_2^T(\xi) \\ &= [(1-\alpha)M_1(\xi^\dagger) + \alpha M_1(\xi^o)] \\ &\quad - [(1-\alpha)M_1(\xi^\dagger) + \alpha M_2(\xi^o)] \\ &\quad \cdot [(1-\alpha)M_3^{-1}(\xi^\dagger) + \alpha M_3^T(\xi^o)]^{-1} \\ &\quad \cdot [(1-\alpha)M_2(\xi^\dagger) + \alpha M_2(\xi^o)]^T \end{aligned}$$

Applying Lemma 4.3.1 gives:

$$M_s(\xi) \geq (1-\alpha)M_1(\xi^\dagger) + \alpha M_1(\xi^o)$$
$$- (1-\alpha)M_2(\xi^\dagger)M_3^{-1}(\xi)M_2^T(\xi)$$
$$- \alpha M_2(\xi^o)M_3^{-1}(\xi^o)M_2^T(\xi^o)$$
$$= (1-\alpha)M_s(\xi^\dagger) + \alpha M_s(\xi^o)$$

Hence, using Lemma 4.3.2:

$$\log \det M_s(\xi)$$
$$\geq \log \det [(1-\alpha)M_s(\xi^\dagger) + \alpha M_s(\xi^o)]$$
$$\geq (1-\alpha) \log \det M_s(\xi^\dagger) + \alpha \log \det M_s(\xi^o)$$

$$\frac{\partial}{\partial\alpha} \log \det M_s(\xi) \Big|_{\alpha=0}$$
$$\geq \log \det M_s(\xi^o) - \log \det M_s(\xi^\dagger)$$
$$\geq 0$$

which violates (2).

This completes the proof of Theorem 4.3.1.

$\nabla\nabla\nabla$

We now extend the equivalence theorem to the case of estimating s out of p parameters.

Theorem 4.3.2 - Equivalence Theorem ($s \leq p$ case).

The following statements are equivalent:

(1) $\xi^\dagger$ maximizes $\det M_s(\xi)$

(2) $\xi^\dagger$ minimizes $\max\limits_{\omega} d_s(\omega,\xi)$

(3) $\max_{\omega} d_s(\omega,\xi) = s$

where: $d_s(\omega,\xi)$ is defined by (4.2.3).

<u>Proof</u>:

We proceed as follows:

(3) $\rightarrow$ (2)

(1) $\rightarrow$ (3)

(3) $\rightarrow$ (1)

(2) $\rightarrow$ (3)

By definition of $d_s(\omega,\xi)$, it follows that:

$$\begin{aligned}
\int_0^{\pi} & d_s(\omega,\xi)\,d\xi(\omega) \\
&= \int_0^{\pi} [h^*(e^{j\omega})\ \overline{M}^{-1}(\xi)\ h(e^{j\omega}) \\
&\quad - h^{*(2)}(e^{j\omega})\ M_3^{-1}(\xi)\ h^{(2)}(e^{j\omega})]\quad d\xi(\omega) \\
&= \mathrm{tr}[\overline{M}^{-1}(\xi)\ \mathit{Re} \int_0^{\pi} h(e^{j\omega})\ h^*(e^{j\omega})\ d\xi(\omega)] \\
&\quad - \mathrm{tr}[M_3^{-1}(\xi)\ \mathit{Re} \int_0^{\pi} h^{(2)}(e^{j\omega})\ h^{*(2)}\ d\xi(\omega)] \\
&= \mathrm{tr}[I_p] - \mathrm{tr}[I_{(p-s)}] = s
\end{aligned}$$

Thus:

$$\max_{\omega} d_s(\omega,\xi) \geq s \tag{4.3.1}$$

Now consider:

$$\max_{\omega} d_s(\omega,\xi^{\dagger}) = \min_{\xi} \max_{\omega} d_s(\omega,\xi) \tag{4.3.2}$$

It follows from (4.3.1) that a sufficient condition for $\xi^{\dagger}$ to satisfy (4.3.2) is:

$$\max_{\omega} d_s(\omega, \xi^{\dagger}) = s$$

Thus: (3) $\rightarrow$ (2)

To show (1) $\rightarrow$ (3):

let $\xi^{o} \in \Xi$ be any design, and consider the design:

$$\xi = (1-\alpha)\xi^{\dagger} + \alpha\xi^{o} \in \Xi$$

Using Theorem 4.3.1 and (4.2.2a), (1) can be written as:

$$\begin{aligned} &\frac{\partial}{\partial\alpha} \log \det M_s[(1-\alpha)\xi^{\dagger} + \alpha\xi^{o}]\Big|_{\alpha} \\ &= \frac{\partial}{\partial\alpha} \log \det \bar{M}[(1-\alpha)\xi^{\dagger} + \alpha\xi^{o}]\Big|_{\alpha=0} \\ &\quad - \frac{\partial}{\partial\alpha} \log \det M_3[(1-\alpha)\xi^{\dagger} + \alpha\xi^{o}]\Big|_{\alpha=0} \\ &\leq 0 \end{aligned} \tag{4.3.3}$$

Now:

$$\begin{aligned} &\frac{\partial}{\partial\alpha} \log \det \bar{M}[(1-\alpha)\xi^{\dagger} + \alpha\xi^{o}]\Big|_{\alpha=0} \\ &= \operatorname{tr} \bar{M}^{-1}(\xi^{\dagger})[\bar{M}(\xi^{o}) - \bar{M}(\xi^{\dagger})] \\ &= \operatorname{tr} \bar{M}^{-1}(\xi^{\dagger})\,\bar{M}(\xi^{o}) - p \end{aligned} \tag{4.3.4}$$

Similarly:

$$\begin{aligned} &\frac{\partial}{\partial\alpha} \log \det M_3[(1-\alpha)\xi^{\dagger} + \alpha\xi^{o}]\Big|_{\alpha=0} \\ &= \operatorname{tr} M_3^{-1}(\xi^{\dagger})\, M_3(\xi^{o}) - (p-s) \end{aligned} \tag{4.3.5}$$

Combining (4.3.3), (4.3.4) and (4.3.5) gives:

$$\frac{\partial}{\partial\alpha} \log \det M_s[(1-\alpha)\xi^{\dagger} + \alpha\xi^{o}]\Big|_{\alpha=0}$$

$$= \operatorname{tr} \bar{M}^{-1}(\xi^{\dagger})\ M(\xi^{o})$$

$$- \operatorname{tr} M_3^{-1}(\xi^{\dagger})\ M_3(\xi^{o}) - s$$

$$\leq 0 \qquad (4.3.6)$$

Let ξ^{o} be a single frequency design at ω, (4.3.6), and the definition of $\bar{M}$ and M_3 for a single frequency design then imply:

$$h^{*}(e^{j\omega})\ \bar{M}^{-1}(\xi^{\dagger})\ h(e^{j\omega})$$
$$- h^{(2)*}(e^{j\omega})\ M_3^{-1}(\xi^{\dagger})\ h^{(2)}(e^{j\omega}) \leq s$$

or:

$$d(\omega,\xi^{\dagger}) - d_{(p-s)}(\omega,\xi^{\dagger}) = d_s(\omega,\xi^{\dagger}) \leq s \qquad (4.3.7)$$

Comparing with (4.3.1) shows that:

(1) $\rightarrow$ (3)

Conversely, if $d_s(\omega,\xi^{\dagger}) = s$ holds, we have (4.3.6) holds for all single frequency design ξ^{o},

hence: (3) $\rightarrow$ (1).

Finally, (4.3.1) and (4.3.7) indicate that $\xi^{\dagger}$ minimizes $\max_{\omega} d_s(\omega,\xi)$,
thus: (2) $\rightarrow$ (3)

This completes the proof of the theorem.

$\nabla\nabla\nabla$

In the next Section, we compare the achievable parameter estimation accuracy for the D_s-optimal and D-optimal designs with a simple example.

4.4 A COMPARISON OF D_s-OPTIMAL AND D-OPTIMAL DESIGNS

In this Section, we compare the D_s-optimal design as defined in Section 4.2, with the D-optimal design.

Consider a first-order autoregressive model.

$$y_k = \frac{b}{1-az^{-1}} u_{k-1} + \frac{1}{1-az^{-1}} e_k$$

the sequences $\{u_k\}$, $\{y_k\}$, $\{e_k\}$ are as defined previously, and the parameter vector

$$\theta^T = (a,b)$$

Following the steps in Chapter 3, we form the average information matrix:

$$\bar{M} = \frac{1}{\sigma^2} \begin{bmatrix} \rho_0 & \frac{1}{b}(\rho_1-\rho_0\alpha_1) \\ \frac{1}{b}(\rho_1-\rho_0\alpha_1) & \frac{1}{b^2}(\rho_0-\sigma^2+\alpha_1^2\rho_0-2\alpha_1\rho_1) \end{bmatrix}$$

and its inverse:

$$\bar{M}^{-1} = \frac{1}{\Delta} \begin{bmatrix} \frac{1}{b^2}(\rho_0-\sigma^2+\alpha_1{}^2\rho_0-2\alpha_1\rho_1) & -\frac{1}{b\sigma^2}(\rho_1-\rho_0\alpha_1) \\ -\frac{1}{b\sigma^2}(\rho_1-\rho_0\alpha_1) & \frac{\rho_0}{\sigma^2} \end{bmatrix}$$

where:

$$\rho_{|j-i|} = E\, y_{k-j}\, y_{k-i}$$

$$E\, e_k^{\,2} = \sigma^2$$

and:

$$\Delta = \frac{1}{b^2\sigma^2}\{\rho_0(\rho_0-\sigma^2)-\rho_1^{\,2}\}$$

Since we are interested in estimating only the paramater a, we formulate the problem for the case p=2, s=1. Following Section 4.2, we arrive at the following criterion for the D_s-optimal design.

$$\max.\ \det M_s = \max.\ \frac{1}{\sigma^2}\{\rho_0 - \frac{(\rho_1-\rho_0\,\alpha_1)^2}{(\rho_0-\sigma^2)+\alpha_1^{\,2}\rho_0-2\alpha_1\rho_1}\}$$

The experimental conditions for the D_s-optimal design are:

$$\rho_1 = \frac{1}{\alpha_1}(\rho_0-\sigma^2) \qquad (4.4.1)$$

The experimental conditions for the D-optimal design (here both parameters are estimated) as given by Chapter 3, are:

$$\rho_1 = 0 \qquad (4.4.2)$$

The variance of the parameter estimate $\hat{a}$ is given by the expression:

$$\text{var } \hat{a} = \sigma^2 \cdot \frac{(\rho_0-\sigma^2)+\alpha_1^2\rho_0-2\alpha_1\rho_1}{\rho_0(\rho_0-\sigma^2)-\rho_1^2} \qquad (4.4.3)$$

A comparison of the two designs, i.e. D_s-optimal and D-optimal designs, is carried out by substituting the respective experimental conditions in Equation (4.4.3). The achievable var $\hat{a}$ for the two designs is shown in the following table:

	var $\hat{a}$
D_s-optimal design	$\sigma^2(\frac{\alpha_1^2}{\rho_0-\sigma^2})$
D-optimal design	$\sigma^2(\frac{1}{\rho_0} + \frac{\alpha_1^2}{\rho_0-\sigma^2})$

From the above table, it is observed that var $\hat{a}$ is lower for the D_s-optimal design.

This example has illustrated that D_s-optimal design is better than D-optimal design in the sense of achievable variance of the parameter estimates.

4.5 A SEQUENTIAL DESIGN ALGORITHM

Sequential design algorithms with proven convergence to a D-optimal design was proposed by Mehra [M4] and extensively studied by Zarrop [Z3]. We shall study, in this Section, an algorithm which is essentially due to Mehra [M4] and Zarrop [Z3], but extends to the D_s-optimality case.

Algorithm

(1) Choose a design $\xi_0 \in \Xi$ such that: $M(\xi_0)$ is non-singular.

(2) Set $k=1$

(3) Choose a frequency ω_k such that:
$$d_s(\omega_k, \xi_k) = \max_{\omega \in [0, \pi]} d_s(\omega, \xi_k)$$

(4) If $d_s(\omega_k, \xi_k) = s$, STOP

(5) Update design to:
$$\xi_{k+1} = (1-\alpha_k)\xi_k + \alpha_k \xi_{\omega_k} \in \Xi$$

(6) Set $k = k+1$; go to (3)

Theorem 4.4.1

If the sequence $\{\alpha_k\}$ in the algorithm is chosen such that:

$$\alpha_k \in [0,1]$$

$$\lim_{k\to\infty} \alpha_k = 0, \qquad \sum_{k=0}^{\infty} \alpha_k = \infty$$

then:

$$\lim_{k\to\infty} \xi_k = \xi^{\dagger} \in \Xi$$

is a D_s-optimal design

Proof:

From (4.2.2a):

$$\frac{\det M_s(\xi_{k+1})}{\det M_s(\xi_k)} = \frac{\det \bar{M}(\xi_{k+1})/\det \bar{M}(\xi_k)}{\det M_3(\xi_{k+1})/\det M_3(\xi_k)} \qquad (4.5.1)$$

Following Zarrop [Z3], we have:

$$M_s(\xi_{k+1}) = (1-\alpha_k)\, M_s(\xi_k) + \alpha_k\, \mathit{Re}\{h(e^{j\omega_k})\; h^*(e^{j\omega_k})\} \qquad (4.5.2)$$

$$\frac{\det M_s(\xi_{k+1})}{\det M_s(\xi_k)} = (1-\alpha_k)^p\, \{1+\beta_k\, d(\omega_k,\xi_k)+\beta_k{}^2 g(\omega_k,\xi_k)\} \qquad (4.5.3)$$

where:

$$\beta_k = \frac{\alpha_k}{1-\alpha_k}$$

$$d(\omega,\xi) = h^*(e^{j\omega})\; \bar{M}^{-1}(\xi)\; h(e^{j\omega})$$

$$g(\omega,\xi) = \tfrac{1}{4}\{d^2(\omega,\xi) - |d_1(\omega,\xi)|^2\}$$

$$d_1(\omega,\xi) = h^T(e^{j\omega})\; \bar{M}^{-1}(\xi)\; h(e^{j\omega})$$

Similarly:

$$M_3(\xi_{k+1}) = (1-\alpha_k)\, M_3(\xi_k)$$

$$+ \alpha_k\, \mathit{Re}\{h^{(2)}(e^{j\omega_k})\; h^{(2)*}(e^{j\omega_k})$$

and:

$$\frac{\det M_3(\xi_{k+1})}{\det M_3(\xi_k)} = (1-\alpha_k)^r \; 1 + \beta_k \; d_r(\omega_k, \xi_k)$$
$$+ \beta_k^2 \; g_r(\omega_k, \xi_k) \qquad (4.5.5)$$

where: $r = p - s$

$$d_r(\omega, \xi) = h^{(2)*}(e^{j\omega}) \; M_3^{-1}(\xi) \; h^{(2)}(e^{j\omega})$$

$$g_r(\omega, \xi) = \tfrac{1}{4}\{d(\omega, \xi) - |d_{1r}(\omega, \xi)|^2\}$$

$$d_{1r}(\omega, \xi) = h^{(2)T}(e^{j\omega}) \; M_3^{-1}(\xi) \; h^{(2)}(e^{j\omega})$$

Combining (4.5.1), (4.5.3) and (4.5.5), gives:

$$\frac{\det M_s(\xi_{k+1})}{\det M_s(\xi_k)} = (1-\alpha)^s \cdot \frac{1+\beta_k d(\omega_k, \xi_k)+\beta_k^2 g(\omega_k, \xi_k)}{1+\beta_k d_r(\omega_k, \xi_k)+\beta_k^2 g_r(\omega_k, \xi_k)} \qquad (4.5.6)$$

Since:

$$\lim_{k\to\infty} \alpha_k = 0$$

the first order expansion of (4.5.6) in α gives:

$$\log \det M_s(\xi_{k+1}) \simeq \log \det M_s(\xi_k)$$
$$- \alpha_k . s + \alpha_k . d - \alpha_k . d_r$$
$$= \log \det M_s(\xi_k) + \alpha_k(d_s - s)$$

Now, $\lim_{k\to\infty} \alpha_k = 0$, therefore there exists k_0, γ dependent on k_0 such that $0 < \gamma < 1$ and for all $k \geq k_0$:

$$\log \det M_s(\xi_{k+1}) > \log \det M_s(\xi_k) + \gamma\alpha_k\{d_s(\omega_k, \xi_k) - s\} \qquad (4.5.7)$$

step (3) of the algorithm ensures that:

$d_s(\omega_k, \xi_k) \geq s$ (cf Theorem 4.3.2) and therefore:

$\{\det M_s(\xi_k),\ k \geq 0\}$

is a monotonically increasing sequence bounded above by $\det M_s(\xi^\dagger)$, therefore:

$$\lim_{k\to\infty} \det M_s(\xi_k) = \det M_s(\xi') \leq \det M_s(\xi^\dagger) \qquad (4.5.8)$$

for some $\xi' \in \Xi$

To show that:

$\det M_s(\xi') = \det M_s(\xi^\dagger)$,

assuming the contrary, then there exists $\varepsilon > 0$ such that:

$d_s(\omega_k, \xi_k) - s \geq \varepsilon$,

for all $k > k_0$

and from (4.5.7):

$$\log \det M_s(\xi_{k_0}) + \varepsilon\gamma \sum_{k=k_0}^{\infty} \alpha_k < \log \det M_s(\xi')$$

By assumption, $\sum_{k=0}^{\infty} \alpha_k = \infty$ and the sequence $\{\det M_s(\xi_k)\}$ is unbounded and this contradicts (4) and thus completes the proof.

$\nabla\nabla\nabla$

4.6 EXAMPLES

The sequential design algorithm presented in the last Section is illustrated by two examples, where: p=2, s=1.

Example 1:

Consider the system (Zarrop [Z3, p.121]):

$$y_k = (b_0+b_1z^{-1})u_{k-1} + \frac{d_0+d_1z^{-1}}{1+c_1z^{-}} e_k$$

Following Section 4.2, we have:

$$\theta^T = (b_0,b_1)$$

$$h_1(z^{-1}) = -\frac{1+c_1z^{-1}}{d_0+d_1z^{-1}} \cdot z^{-1}$$

$$h_2(z^{-1}) = -\frac{1+c_1z^{-1}}{d_0+d_1z^{-1}} \cdot z^{-2}$$

$$\bar{M} = \frac{1+c_1{}^2+2c_1\cos\omega}{d_0{}^2+d_1{}^2+2d_0d_1\cos\omega}\begin{bmatrix} 1 & \cos\omega \\ \cos\omega & 1 \end{bmatrix}$$

Estimating b_0 using the sequential design algorithm given in Section 4.5, where $\alpha_k = \frac{1}{k+1}$ is used, the following table gives the final design for several values of c_1, d_0 and d_1 with initial design $\omega = 0.5$, $\lambda = 1$.

			Final Design	
c_1	d_0	d_1	ω	λ
.5	1	.3	1.46	1
.1	-.8	.3	1.11	1
-.5	-1	-.3	2.10	1

In all cases, step (3) of the algorithm chooses the optimal frequency after several iterations. The convergence characteristics for three different initial designs for the case, $c_1 = .5$, $d_0 = 1$, $d_1 = .3$, are shown in Figure 4.1.

It is interesting to compare the D_s-optimal and D-optimal designs for this system.

For $c_1 = .5$, $d_0 = 1$, $d_1 = .3$, the D-optimal design gives (Zarrop [Z3, p.121]):

$$\omega = 1.3804$$

and $$\text{var } \hat{b}_0 = 0.8673$$

The D_s-optimal design gives:

$$\omega = 1.46$$

and $$\text{var } \hat{b}_0 = 0.8604$$

It can be seen from these results that the D_s-optimal design achieves a lower value on the variance of the parameter estimate.

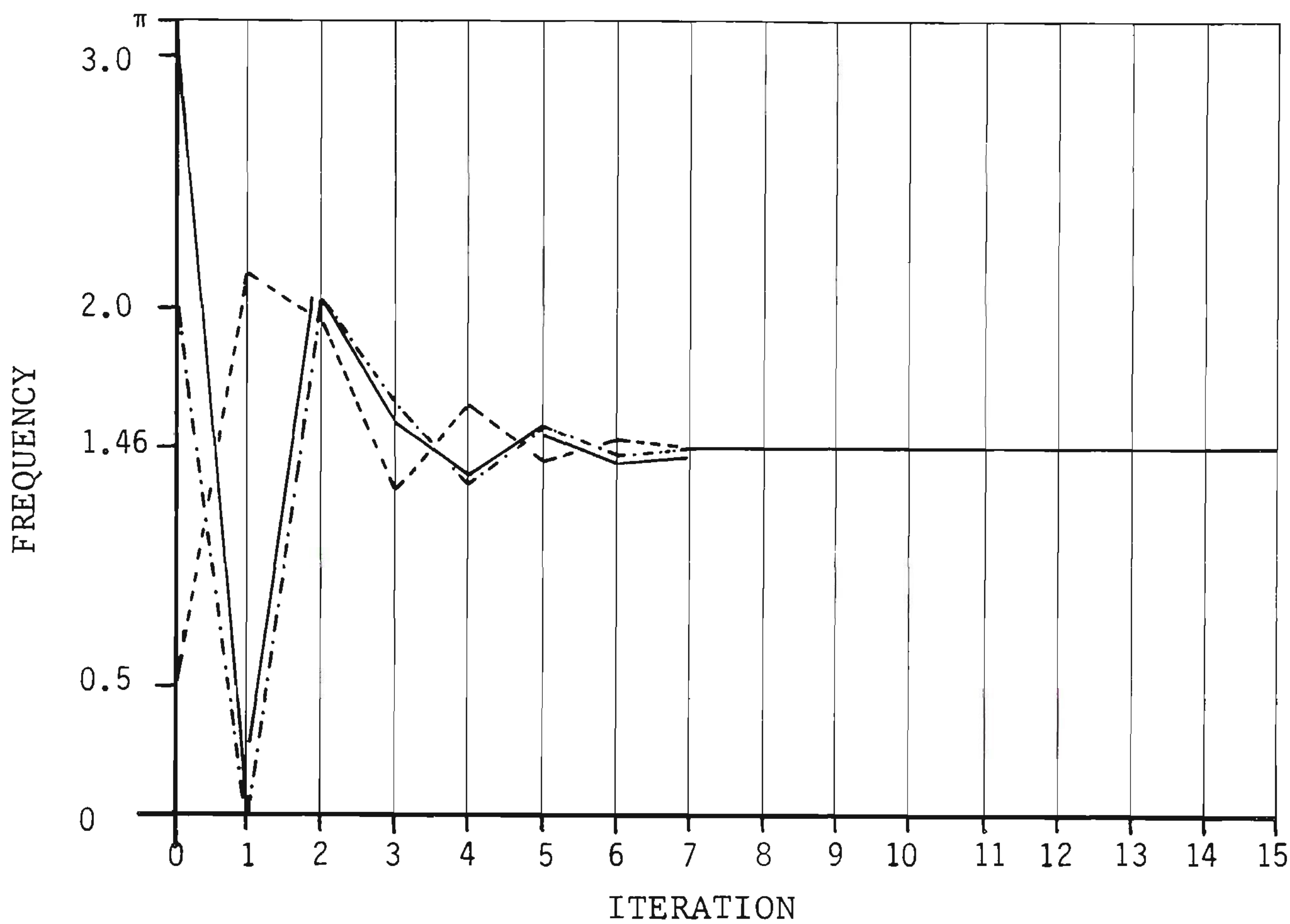


Fig. 4.1: Convergence characteristics for different initial design frequencies for the case $c_1=0.5$, $d_0=1.0$, $d_1=0.3$

<u>Example 2</u>:

Consider the system:

$$y_k = \frac{b_0}{1+a_1 z^{-1}} u_{k-1} + \frac{d_0+d_1 z^{-1}}{1+c_1 z^{-1}} e_k$$

Following Section 4.2, we have:

$$\theta^T = (a_1, b_0)$$

$$h_1(z^{-1}) = b_0 \frac{1+c_1 z^{-1}}{d_0+d_1 z^{-1}} \frac{z^{-1}}{(1+a_1 z^{-1})^2}$$

$$h_2(z^{-1}) = - \frac{1+c_1 z^{-1}}{d_0+d_1 z^{-1}} \frac{z^{-1}}{1+a_1 z^{-1}}$$

and

$$\bar{M} = b_0^2 \frac{1+c_1^2+2c_1\cos\omega}{d_0^2+d_1^2+2d_0 d_1\cos\omega} \frac{1}{(1+a_1^2+2a_1\cos\omega)^2}$$

$$\cdot \begin{bmatrix} 1 & -\frac{1}{b_0}(a_1+\cos\omega) \\ -\frac{1}{b_0}(a_1+\cos\omega) & \frac{1}{b_0^2}(1+a_1^2+2a_1\cos\omega) \end{bmatrix}$$

Estimating a_1 using the sequential design algorithm in Section 4.5 with $\alpha_k = \frac{1}{k+1}$, the final design for several values of a_1, b_0, c_1, d_0, d_1 are tabulated below.

					Final Design			
a	b_0	c_1	d_0	d_1	ω_1	λ_1	ω_2	λ_2
-0.1	1.0	-0.1	1.0	0.0	0	0.5	π	0.5
-0.3	0.5	0.5	1.0	0.7	0	0.5	π	0.5

In all cases, the initial design is $\omega = 0.5$, $\lambda = 1$. The same final design is also obtained with different initial frequencies. In all cases tried, the final design always converges to two frequencies.

4.7 CONCLUSION

Optimal input design for identification experiments based on two different criteria has been shown to be equivalent for the case of estimating part of the system parameters. A proven sequential design algorithm which converges to a D-optimal design has also been extended for the D_s-optimality case, and illustrated by simple examples. A comparison of D_s-optimal and D-optimal designs has been carried out for the case s=1, p=2.

The D_s-optimal design was based on the assumption that $\bar{M}(\xi)$ is non-singular. Methods used to overcome singular $\bar{M}(\xi)$ in linear regression problems, e.g. transformations used in Kiefer [K4] and Atwood [A4] cannot be extended to the dynamic case.

APPENDIX 4.1 EXAMPLE

Another example to illustrate the sequential design algorithm of Section 4.5 is given below. Here, s=1, p=2.

Consider the system:

$$y_k = \frac{1}{1+a_1 z^{-1}+a_2 z^{-2}} u_{k-1} + \frac{d_0+d_1 z^{-1}}{1+c_1 z^{-1}} e_k$$

Following Section 4.2, we have:

$$\theta^T = (a_1, a_2)$$

$$h_1(z^{-1}) = \frac{1+c_1 z^{-1}}{d_0+d_1 z^{-1}} \cdot \frac{1}{(1+a_1 z^{-1}+a_2 z^{-2})^2} \cdot z^{-2}$$

$$h_2(z^{-1}) = \frac{1+c_1 z^{-1}}{d_0+d_1 z^{-1}} \cdot \frac{1}{(1+a_1 z^{-1}+a_2 z^{-2})^2} \cdot z^{-3}$$

and:

$$\bar{M} = \frac{(1+c_1 z^{-1})(1+c_1 z)}{(d_0+d_1 z^{-1})(d_0+d_1 z)} \quad \frac{1}{(1+a_1 z^{-1}+a_2 z^{-2})^2(1+a_1 z+a_2 z^{-2})^2}$$

$$\cdot \begin{bmatrix} 1 & \cos\omega \\ \cos\omega & 1 \end{bmatrix}$$

Estimating a_1 using the sequential design algorithm in Section 4.5 with $\alpha_k = \frac{1}{k+1}$, the following table gives the final design for several values of a_1, a_2, c_1, d_0, d_1 with initial design $\omega = 0.5$, $\lambda = 1$.

					Final Design	
a_1	a_2	c_1	d_0	d_1	ω	λ
1.0	0.25	-0.1	1.0	0.0	1.98	1
0.1	-0.30	0.5	-0.1	-0.6	2.78	1
1.3	0.90	-0.7	0.6	0.2	1.91	1

The same final design is also obtained with different initial frequencies. In all cases tried, the final design always converges to a single frequency after a number of iterations.

CHAPTER 5

EFFICIENCY OF MINIMAL UNIFORM DESIGNS

CHAPTER 5: EFFICIENCY OF MINIMAL UNIFORM DESIGNS

5.1 INTRODUCTION

In this Chapter we look at the problem of comparing minimal uniform designs with a D-optimal design.

D-efficiency is defined in Section 5.2, and the efficiency of a D-optimal design is considered equal to unity. In Section 5.3, we consider some examples to show the efficiency of minimal uniform designs, and in Section 5.4, we carry out a detailed analysis of a two-parameter system.

In the Appendix, general expressions for $|G_i||G_i^*|$, $i=1,\ldots,\binom{2\ell}{p}$ for a general system with p=2, are given.

5.2 PROBLEM STATEMENT

Consider a linear, time invarient, stable, SISO, discrete time system described by (2.2.1).

Following the derivation in Chapter 2, the average information matrix for the system parameters is given by:

$$\bar{M}\ (\xi) \quad = \quad Re \int_{\Omega} h(e^{j\omega})\ h^{*}(e^{j\omega})\ d\xi(\omega) \qquad (5.2.1)$$

where: the total number of system parameters $p = m+n+1$;

Ω denotes the set: $\Omega \quad = \quad \{\omega/0 < \omega < \pi\}$

ξ is a normalized design and $h_i(z^{-1})$ is given in Chapter 2 (Sec.2.3).

A design containing a pure discrete spectrum can be characterized by:

$$\{\omega_1, \ldots,\ \omega_\ell :\ \ \lambda_1, \ldots,\ \lambda_\ell\}$$

where: ω_1, $i=1, \ldots, \ell$ represents frequency, and the positive weights $\lambda_\ell, \ldots, \lambda_\ell$ sum to unity.

With a pure discrete spectrum, (5.2.1) can be written as:

$$\bar{M}\ (\xi) \quad = \quad G\ \Lambda\ G^{*} \qquad (5.2.2)$$

where: $G \quad = \quad [h(\omega_1)\ \bar{h}(\omega_1)\ h(\omega_2)\ \bar{h}(\omega_2)\ \ldots,\ h(\omega_\ell)\ \bar{h}(\omega_\ell)]$

(5.2.2a)

$$\Lambda = \mathrm{diag}\ [\ \frac{\lambda_1}{2}\ \frac{\lambda_1}{2}\ \frac{\lambda_2}{2}\ \frac{\lambda_2}{2}\ \ldots,\ \frac{\lambda_\ell}{2}\ \frac{\lambda_\ell}{2}\ ]$$

and $\quad G^* = (\bar{G})^T$

Forming the determinant and using a matrix property (Fedorov [F1]):

$$|\bar{M}\ (\xi)| = \sum_i |G_i|\,|G_i^{\ *}| \prod_{i\alpha=1}^{p} \Lambda_{i\alpha} \qquad (5.2.3)$$

where G_i is a p x p minor of G, and the sum extends over all $\binom{2\ell}{p}$ possible minors composed of p columns numbered $i_1, i_2, \ldots, i_p$.

To compare the efficiency of two designs, we define a measure of efficiency as follows:

$$\eta = \left[\frac{|\bar{M}\ (\xi)|}{|\bar{M}\ (\xi')|}\right]^{1/p} \qquad (5.2.4)$$

The pth root of the quantity $|\bar{M}\ (\xi)|/|\bar{M}\ (\xi')|$ eliminates the dependence of efficiency on the dimension of the parameter vector. This definition can be interpreted as comparing the average estimation error per parameter for the two designs. For $\xi' = \xi^\dagger$, a D-optimal design, we call η the D-efficiency of the design ξ.

5.3 EFFICIENCY OF MINIMAL UNIFORM DESIGNS

We consider a uniform design with the minimum number of input frequencies $\omega_i \in (0,\pi)$, i=1,2..., γ where:

$$\gamma = p/2 \quad \text{for p even}$$
$$= (p+1)/2 \quad \text{for p odd}$$

A uniform design is one which has equal weights on all input frequencies. Thus the search for an optimal input reduces from a 2p dimensional space to a p/2 or (p+1)/2 dimensional space.

Denote:

$$\beta_i = \prod_{i\alpha=1}^{p} \Lambda_{i\alpha}$$

then for a ℓ frequency, D-optimal design (5.2.3) becomes:

$$|\bar{M}(\xi^\dagger)| = \sum_{i=1}^{\Gamma} \beta_i \, |G_i|\,|G_i^*| \tag{5.3.1}$$

where: $\Gamma = \binom{2\ell}{p}$

Define:

(i) $\xi^\dagger = \{\omega_1,\ldots,\omega_\ell;\ \lambda_1,\ldots,\lambda_\ell\}$
a D-optimal design, where $\gamma \leq \ell \leq p$

(ii) $\zeta = \{\omega_1,\ \omega_2,\ldots,\omega_\gamma\ ;\ \lambda_1,\ldots,\lambda_\gamma\}$
a γ frequency uniform design,
i.e. $\lambda_1 = \lambda_2 = \ldots, = \lambda_\gamma$

(iii) $\xi' = \{\omega_1, \ldots, \omega_\gamma; \quad \lambda_1, \ldots, \lambda_\gamma / \omega_i \in \Omega^\ell\}$
a γ frequency uniform design, where Ω^ℓ denotes the set of ℓ frequencies in the D-optimal design $\xi^\dagger$.

Note that:

(i) $$|\bar{M}(\xi')| = p^{-p} |G(p)||G^*(p)| \qquad (5.3.2)$$

where:

$$|G(p)|\cdot|G^*(p)| = \max_j \{|G_j|\cdot|G_j^*|, \; j=1,\ldots,\tbinom{\ell}{\gamma}\} \qquad (5.3.2a)$$

and G_j has the form:

$$[h(\omega_{i_1})\,\bar{h}(\omega_{i_1}), \; \ldots, \; h(\omega_{i_\gamma})\,\bar{h}(\omega_{i_\gamma})] \qquad (5.3.2b)$$

(ii) $$|\bar{M}(\zeta)| \geq |\bar{M}(\xi')|$$

Then, from the definition of efficiency (5.2.4), it follows that:

$$\eta = \left[\frac{|\bar{M}(\zeta)|}{|\bar{M}(\xi^\dagger)|}\right]^{1/p}$$

$$\geq \left[\frac{|\bar{M}(\xi')|}{|\bar{M}(\xi^\dagger)|}\right]^{1/p}$$

$$= \left[\frac{p^{-p}\;|G(p)|\cdot|G^*(p)|}{\sum_i \beta_i\;|G_i||G_i^*|}\right]^{1/p} \qquad (5.3.3)$$

$$\geq \left[\frac{p^{-p} \; |G(p)| \, |G^{*}(p)|}{\max\limits_{i} \{|G_i| \, |G_i^{*}|\} \sum\limits_{i} \beta_i} \right]^{1/p} \qquad (5.3.3a)$$

the r.h.s. of the inequality (5.3.3a) represents the achievable lower bound for D-efficiency of a minimal uniform design.

Now, we illustrate the efficiency of minimal uniform designs by giving a number of examples:

Example 1

Consider the following system (Zarrop [Z3, p.121]):

$$y_k = (b_0 + b_1 z^{-1}) \; u_k + \frac{1}{1 + c_1 z^{-4}} \; e_k \qquad (5.3.4)$$

The information matrix for the parameter vector:

$$\theta = (b_0, b_1)^T$$

is given by:

$$\bar{M}(\xi) = \sum_{i=1}^{\ell} \lambda_i \; f(\omega_i) \begin{bmatrix} 1 & \cos\omega_i \\ \cos\omega_i & 1 \end{bmatrix}$$

where: $f(\omega) = 1 + c_1^{\,2} + 2c_1 \cos 4\omega$

Note: $\bar{M}(\xi)$ is independent of b_0 and b_1.

Then the D-optimal design is (Zarrop [Z3]):

$$\omega_1 = \pi/4 \; , \; \omega_2 = 3\pi/4$$

$$\lambda_1 = \lambda_2 = \tfrac{1}{2}$$

which yields (for $c_1 = -0.06$):

$$|\bar{M}\ (\xi^{\dagger})| = 1.2625$$

Substituting the values of ω_1 and ω_2 in (A5.1.13), we get:

$$|G_i||G_i^*| = 2.5250\ ,\ i=1,\ldots,4$$

$$= 5.0499\ ,\ i=5,6$$

Then, from (5.3.2a), we have $|G(p)||G^*(p)| = 2.5250$ and the determinant of the p/2 frequency uniform design ξ' given by (5.3.2) is:

$$|\bar{M}\ (\xi') = p^{-p}\ |G(p)||G^*(p)|$$

$$= 2^{-2}\ (2.5250)$$

$$= 0.6312$$

The efficiency for the design ξ' is then given by:

$$\eta = \left[\frac{|\bar{M}\ (\xi')|}{|\bar{M}\ (\xi^{\dagger})|}\right]^{\frac{1}{2}}$$

$$= \{\tfrac{1}{2}\}^{\frac{1}{2}}$$

$$= 0.707$$

Example 2

Consider the system:

$$y_k = (b_0+b_1 z^{-1})\ u_k + (1+d_1 z^{-4})\ e_k \qquad (5.3.5)$$

where: $\theta = (b_0, b_1)^T$

the information matrix is given by:

$$\bar{M}\ (\xi) = \sum_{i=1}^{\ell} \lambda_i\ f(\omega_i) \begin{bmatrix} 1 & \cos\omega_i \\ \cos\omega_i & 1 \end{bmatrix}$$

where: $f(\omega) = (1+d_1{}^2+2d_1\ \cos 4\omega)^{-1}$

then the D-optimal design is computed as:

$$\omega_1 = \pi/4\ ,\ \omega_2 = 3\pi/4$$

$$\lambda_1 = \lambda_2 = \tfrac{1}{2}$$

which yields (for $d_1 = 0.1$):

$$|\bar{M}\ (\xi^{\dagger})| = 1.5242$$

Substituting the values of ω_1 and ω_2 in (A5.1.13), and after simplification, we get:

$$|G_i||G_i^*| = 3.0483\ ,\ i=1,\ldots,4$$

$$= 6.0966\ ,\ i=5,6$$

Then, from (5.3.2a) we get $|G(p)||G^*(p)| = 3{,}0483$ and the determinant of the design ξ' given by (5.3.2) is:

$$|\bar{M}(\xi')| = p^{-p}\,|G(p)|\,|G^{*}(p)|$$

$$= 2^{-2}(3.0483)$$

$$= 0.7621$$

The efficiency for ξ' is then given by:

$$\eta = \left[\frac{|\bar{M}(\xi')|}{|\bar{M}(\xi^{\dagger})|}\right]^{\frac{1}{2}}$$

$$= \{\tfrac{1}{2}\}^{\frac{1}{2}}$$

$$= 0.707$$

5.4 ANALYSIS OF EFFICIENCY OF A TWO-PARAMETER SYSTEM

In this Section we give an analysis of the D-efficiency of a minimal uniform input design, for a two-parameter system.

Consider the model given by (5.3.4), and solving for $|G_i||G_i^*|$ using the expressions (A5.1.13), we arrive at the following:

$$|G_1||G_1^*| = 4f^2(\omega_1)\ \{1-\cos^2\omega_1\} \tag{5.4.1}$$

$$|G_2||G_2^*| = 4f^2(\omega_2)\ \{1-\cos^2\omega_2\} \tag{5.4.2}$$

$$|G_i||G_i^*| = 2f(\omega_1)\ f(\omega_2)\ \{1-\cos(\omega_1-\omega_2)\},\ i=3,4 \tag{5.4.3}$$

$$|G_i||G_i^*| = 2f(\omega_1)\ f(\omega_2)\ \{1-\cos(\omega_1+\omega_2)\},\ i=5,6 \tag{5.4.4}$$

where: $f(\omega) = 1 + c_1{}^2 + 2c_1 \cos 4\omega$ and $-1 < c_1 < 0$

The optimal frequencies are $\omega_1 = \pi/4$, $\omega_2 = 3\pi/4$. It is noted that:

$$f(\omega) = f(\omega_1) = f(\omega_2),$$

and substituting the values of ω_1, ω_2 in (5.4.1) to (5.4.4), yields the following simplified expressions for $|G_i||G_i^*|$:

$$|G_i||G_i^*| = 2f^2(\omega)\ ,\ i=1,\ldots,4$$

$$= 4f^2(\omega)\ ,\ i=5,6$$

Now, substituting the values of $|G_i||G_i^*|$ and β_i in (5.3.1) gives:

$$|\bar{M}\ (\xi)| = f^2(\omega)$$

where $\beta_i = 1/16$, since $\lambda_1 = \lambda_1 = \frac{1}{2}$

Now, from the definition (5.3.2a) we get:

$$|G(p)| \cdot |G^*(p)| = 2f^2(\omega)$$

and (5.3.2) gives:

$$|\bar{M}\ (\xi')| = p^{-p}\ |G(p)||G^*(p)|$$

$$= 2^{-2}\ \{2f^2(\omega)\}$$

$$= \tfrac{1}{2}\ f^2(\omega)$$

where ξ' is the minimal uniform design defined in Section 5.3.

The D-efficiency of design ξ' is given by:

$$\eta = \left[\frac{|\bar{M}\ (\xi')|}{|\bar{M}\ (\xi^{\dagger})|}\right]^{1/p}$$

$$= \left[\frac{\tfrac{1}{2}\ f^2(\omega)}{f^2(\omega)}\right]^{\frac{1}{2}}$$

$$= \{\tfrac{1}{2}\}^{\frac{1}{2}}$$

$$= 0.707$$

<u>Remark</u>

(i) The D-efficiency of the uniform minimal design ξ' for the system (5.3.4) is always equal to 0.707, where $-1 < c_1 < 0$.

(ii) The efficiency of the uniform minimal design ζ given by:

$$\eta' = \left[\frac{|\bar{M}(\zeta)|}{|\bar{M}(\xi^{\dagger})|} \right]^{1/p}$$

for the system (5.3.4), varies with the value of c_1; this is illustrated in Figure 5.1.

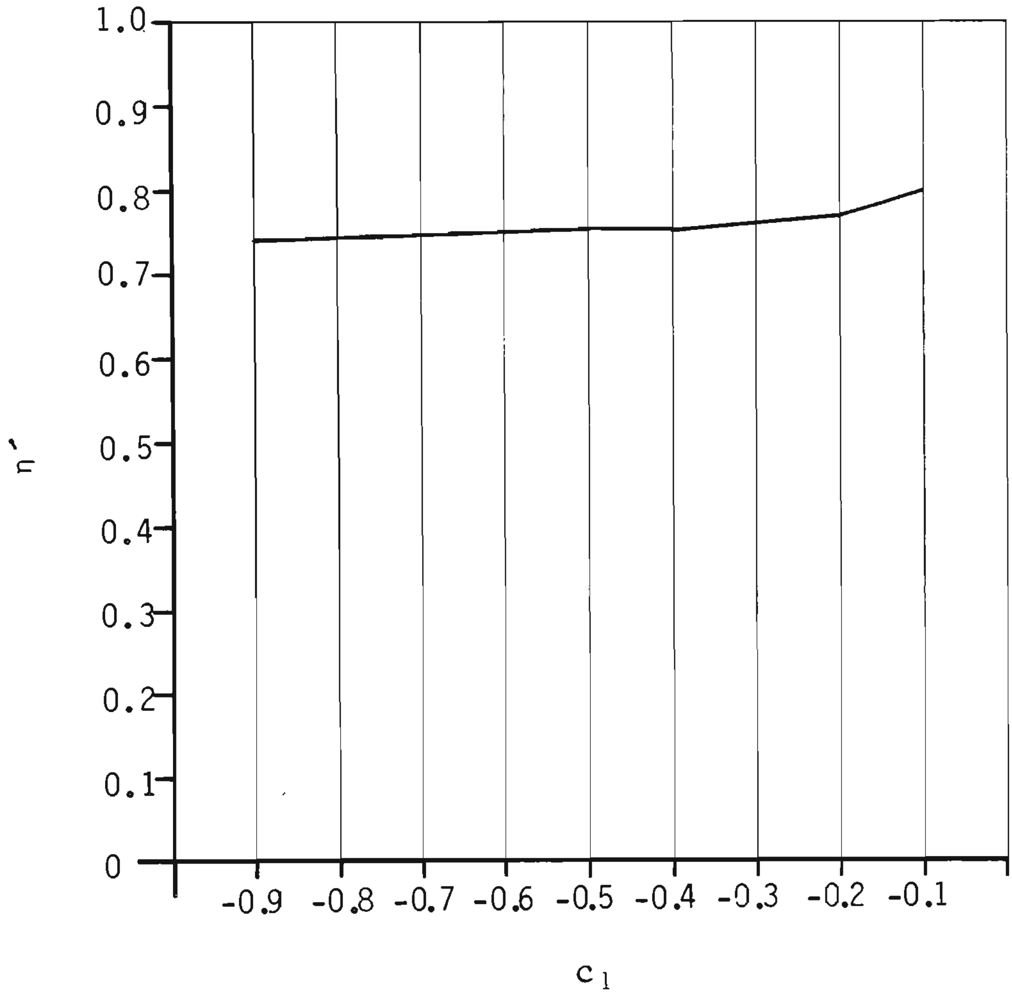


Fig. 5.1: Variation of η' with parameter c_1

Now, consider the two-parameter system given by (5.3.5). The expressions for $|G_i||G_i^*|$ are similar to (5.4.1) to (5.4.2), where $f(\omega) = (1+d_1{}^2+2d_1\cos4\omega)^{-1}$ and $0 < d_1 < 1$.

Substituting the optimal frequencies $\omega_1 = \pi/4$, $\omega_2 = 3\pi/4$ in $f(\omega)$, we observe that:

$$f(\omega) = f(\omega_1) = f(\omega_2)$$

Then the expressions for $|G_i||G_i^*|$, after simplification, give:

$$\begin{aligned} |G_i||G_i^*| &= 2f^2(\omega) \;,\; i=1,\ldots,4 \\ &= 4f^2(\omega) \;,\; i=5,6 \end{aligned}$$

Next, substituting the values of $|G_i||G_i^*|$ and β_i in (5.3.1) yields:

$$|\bar{M}\,(\xi^{\dagger})| = f^2(\omega)$$

and from (5.3.2a) we get:

$$|G(p)||G^*(p)| = 2f^2(\omega)$$

The D-efficiency of design ξ' is then:

$$\begin{aligned} \eta &= \{\tfrac{1}{2}\}^{\frac{1}{2}} \\ &= 0.707 \end{aligned}$$

Remark

(i) The efficiency of the uniform minimal design ξ' is always equal to 0.707, where $0 < \mathbf{d}_1 < 1$.

(ii) The efficiency of the uniform minimal design ζ given by:

$$\eta' = \left[\frac{|\bar{M}\ (\zeta)|}{|\bar{M}\ (\xi^{\dagger})|} \right]^{\frac{1}{2}}$$

varies with the value of d_1, as shown in Figure 5.2.

(iii) It is observed from Figure 5.2 that the efficiency of design ζ is $\eta' = 0.707$ when $d_1 = 0.8$. This value of efficiency $\eta' = 0.707$ is equal to the efficiency of the minimal uniform design ξ'.

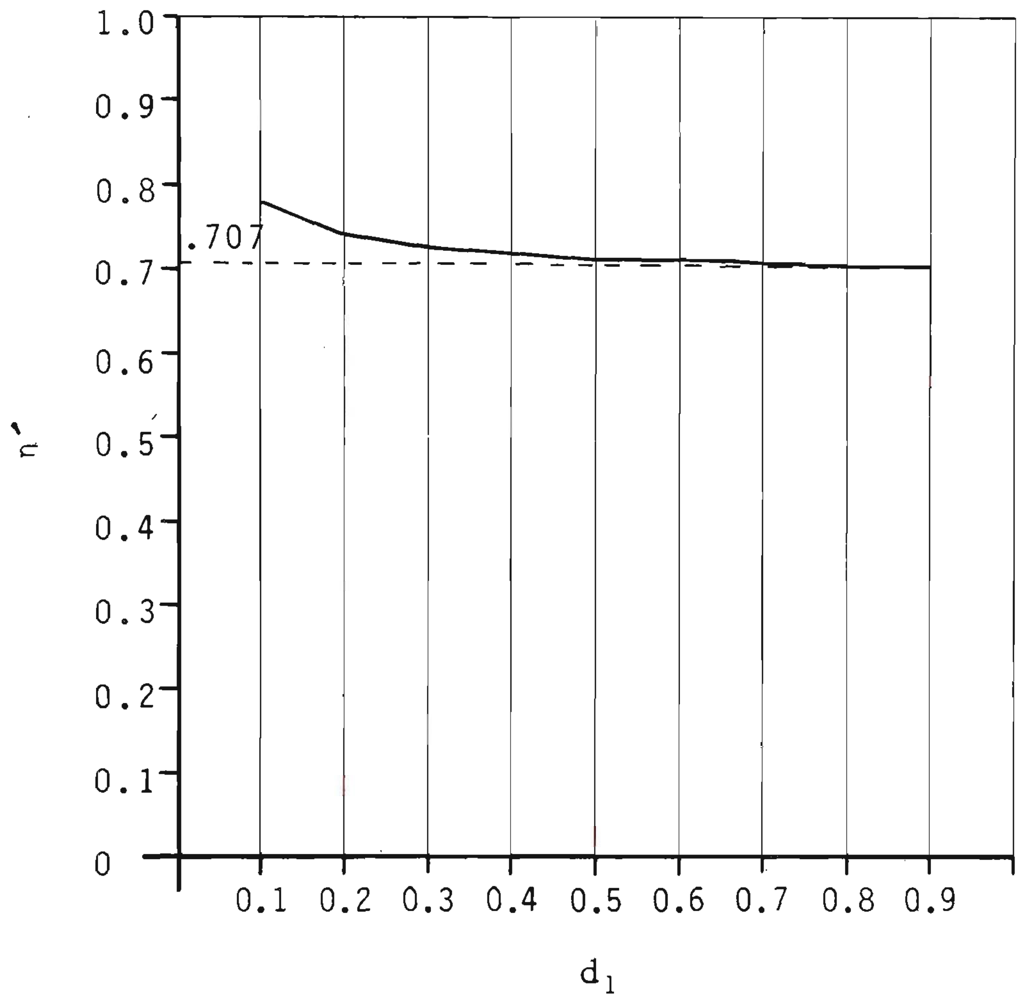


Fig. 5.2: Variation of η' with parameter d_1

Final Remark

For the systems (5.3.4) or (5.3.5), the efficiency is $\eta = 0.707$. Using (5.3.3a), the possible achievable lower bound is 0.6.

Thus, for the system (5.3.4) or (5.3.5), the D-efficiency of the minimal uniform design ξ' does not achieve the lower bound.

5.5 CONCLUSION

The problem of comparing the efficiency of minimal uniform designs has been considered. Some first order systems were considered to illustrate the D-efficiency for minimal uniform designs. It was shown in Section 5.4 that, for the models (5.3.4) and (5.3.5), the efficiency of minimal uniform design ξ' is 0.707; whereas the efficiency for minimal uniform design ζ varies with the value of parameters c_1 and d_1 for models (5.3.4) and (5.3.5) respectively.

Some difficulties were encountered in generalizing the lower bound of D-efficiency given by (5.3.3a). The main problem was to get a relationship between $|G(p)||G^*(p)|$ and $|G_j||G_j^*|$. When G_j has the same form given by (5.3.2b), then from definition (5.3.2a):

$$|G(p)||G^*(p)| \geq |G_j||G_j^*|$$

But when G_j is not of the form of (5.3.2b), then the above relationship does not hold.

APPENDIX 5.1: EXPRESSIONS FOR $|G_i||G_i^*|$

Consider a general model described by (2.2.1) with p=2 and $p/2 \leq \ell \leq p$; in general $\ell = p$.

We proceed to derive general expressions for $|G_i||G_i^*|$, $i=1,\ldots,\binom{2\ell}{p}$ where G_i is a p x p minor of G; G_i and G are described in Section 5.2.

It follows from (5.3.2a), that for the case p=2, $\ell=2$:

$$G = [h(\omega_1)\ \bar{h}(\omega_1)\ h(\omega_2)\ \bar{h}(\omega_2)]$$

and $$\Lambda = \text{diag}\left[\frac{\lambda_1}{2}\ \frac{\lambda_1}{2}\ \frac{\lambda_2}{2}\ \frac{\lambda_2}{2}\right]$$

where: $$h(\omega) = \begin{bmatrix} h_1(\omega) \\ h_2(\omega) \end{bmatrix}$$

and $$\binom{2\ell}{p} = 6$$

We now proceed to form the minors G_i, $i=1,\ldots,6$ and the corresponding G_i^*'s.

Forming the $\binom{2\ell}{p}$ minors of G, as follows:

$$G_1 = [h(\omega_1)\ \bar{h}(\omega_1)] = \begin{bmatrix} h_1(\omega_1) & \bar{h}_1(\omega_1) \\ h_2(\omega_1) & \bar{h}_2(\omega_1) \end{bmatrix} \tag{A5.1.1}$$

$$G_2 = [h(\omega_2)\ \bar{h}(\omega_2)] = \begin{bmatrix} h_1(\omega_2) & \bar{h}_1(\omega_2) \\ h_2(\omega_2) & \bar{h}_2(\omega_2) \end{bmatrix} \tag{A5.1.2}$$

$$G_3 = [h(\omega_1)\ h(\omega_2)] = \begin{bmatrix} h_1(\omega_1) & h_1(\omega_2) \\ h_2(\omega_1) & h_2(\omega_2) \end{bmatrix} \tag{A5.1.3}$$

$$G_4 = [\bar{h}(\omega_1)\ \bar{h}(\omega_2)] = \begin{bmatrix} \bar{h}_1(\omega_1) & \bar{h}_1(\omega_2) \\ \bar{h}_2(\omega_1) & \bar{h}_2(\omega_2) \end{bmatrix} \tag{A5.1.4}$$

$$G_5 = [h(\omega_1)\ \bar{h}(\omega_2)] = \begin{bmatrix} h_1(\omega_1) & \bar{h}_1(\omega_1) \\ h_2(\omega_1) & \bar{h}_2(\omega_2) \end{bmatrix} \tag{A5.1.5}$$

$$G_6 = [\bar{h}(\omega_1)\ h(\omega_2)] = \begin{bmatrix} \bar{h}_1(\omega_1) & h_1(\omega_2) \\ \bar{h}_2(\omega_1) & h_2(\omega_2) \end{bmatrix} \tag{A5.1.6}$$

The corresponding G_i^*, i=1,...,6, are:

$$G_1^* = \begin{bmatrix} \bar{h}^T(\omega_1) \\ h^T(\omega_1) \end{bmatrix} = \begin{bmatrix} \bar{h}_1(\omega_1) & \bar{h}_2(\omega_1) \\ h_1(\omega_1) & h_2(\omega_1) \end{bmatrix} \quad (A5.1.7)$$

$$G_2^* = \begin{bmatrix} \bar{h}^T(\omega_2) \\ h^T(\omega_2) \end{bmatrix} = \begin{bmatrix} \bar{h}_1(\omega_2) & \bar{h}_2(\omega_2) \\ h_1(\omega_2) & h_2(\omega_2) \end{bmatrix} \quad (A5.1.8)$$

$$G_3^* = \begin{bmatrix} \bar{h}^T(\omega_1) \\ h^T(\omega_2) \end{bmatrix} = \begin{bmatrix} \bar{h}_1(\omega_1) & \bar{h}_2(\omega_1) \\ h_1(\omega_2) & h_2(\omega_2) \end{bmatrix} \quad (A5.1.9)$$

$$G_4^* = \begin{bmatrix} h^T(\omega_1) \\ h^T(\omega_2) \end{bmatrix} = \begin{bmatrix} h_1(\omega_1) & h_2(\omega_1) \\ h_1(\omega_2) & h_2(\omega_2) \end{bmatrix} \quad (A5.1.10)$$

$$G_5^* = \begin{bmatrix} \bar{h}^T(\omega_1) \\ h^T(\omega_2) \end{bmatrix} = \begin{bmatrix} \bar{h}_1(\omega_1) & \bar{h}_2(\omega_1) \\ h_1(\omega_2) & h_2(\omega_2) \end{bmatrix} \quad (A5.1.11)$$

$$G_6^* = \begin{bmatrix} h^T(\omega_1) \\ \bar{h}^T(\omega_2) \end{bmatrix} = \begin{bmatrix} h_1(\omega_1) & h_2(\omega_1) \\ \bar{h}_1(\omega_2) & \bar{h}_2(\omega_2) \end{bmatrix} \quad (A5.1.12)$$

Taking determinants on both sides of (A5.1.1) and (A5.1.2), we get:

$$|G_1| = h_1(\omega_2)\ \bar{h}_2(\omega_1) - \bar{h}_1(\omega_1)\ h_2(\omega_1)$$

$$|G_1^*| = \bar{h}_1(\omega_1)\ h_2(\omega_1) - h_1(\omega_1)\ \bar{h}_2(\omega_1)$$

Taking the product of $|G_1|$ and $|G_1^*|$, and after simplification, we arrive at the following expression for $|G_1||G_1^*|$:

$$|G_1||G_1^*| = 2\ \{h_1(\omega_1)\ \bar{h}_1(\omega_1)\ h_2(\omega_1)\ \bar{h}_2(\omega_1)\}$$

$$- \{h_1(\omega_1)\ \bar{h}_2(\omega_1)\}^2 - \{h_2(\omega_1)\ \bar{h}_1(\omega_1)\}^2$$

Similarly, the expressions for $|G_i||G_i^*|$, i=2,...,6 are obtained by taking the product of $|G_i|$ and $|G_i^*|$. After simplification, the final expressions for $|G_i||G_i^*|$ are:

$$|G_i||G_i^*| = 2\ \{h_1(\omega_i)\ \bar{h}_1(\omega_i)\ h_2(\omega_i)\ \bar{h}_2(\omega_i)\}$$

$$- \{h_1(\omega_i)\ \bar{h}_2(\omega_i)\}^2 - \{h_2(\omega_i)\ \bar{h}_1(\omega_i)\}^2$$

$$i=1,2$$

$$= h_1(\omega_1)\ \bar{h}_1(\omega_1)\ h_2(\omega_2)\ \bar{h}_2(\omega_2)$$

$$+ h_1(\omega_2)\ \bar{h}_1(\omega_2)\ h_2(\omega_1)\ \bar{h}_2(\omega_1)$$

$$- h_1(\omega_1)\ \bar{h}_2(\omega_1)\ h_2(\omega_2)\ \bar{h}_1(\omega_2)$$

$$- h_1(\omega_2)\ \bar{h}_2(\omega_2)\ h_2(\omega_1)\ \bar{h}_1(\omega_1)$$

$$i=3,4$$

$$
\begin{aligned}
= \; & h_1(\omega_1)\,\bar{h}_1(\omega_1)\,h_2(\omega_2)\,\bar{h}_2(\omega_2) \\
& + h_1(\omega_2)\,\bar{h}_1(\omega_2)\,h_2(\omega_1)\,\bar{h}_2(\omega_1) \\
& - h_1(\omega_1)\,\bar{h}_2(\omega_1)\,h_1(\omega_2)\,\bar{h}_2(\omega_2) \\
& - h_2(\omega_2)\,\bar{h}_1(\omega_2)\,\bar{h}_1(\omega_1)\,h_2(\omega_1)
\end{aligned}
$$

$$i=5,6 \qquad \text{(A5.1.13)}$$

Note that:

(i) $G_4 = \bar{G}_3$

$G_6 = \bar{G}_5$

(ii) $|G_4|\,|G_4^*| = |G_3|\,|G_3^*|$

$|G_6|\,|G_6^*| = |G_5|\,|G_5^*|$

CHAPTER 6

OPTIMUM EXPERIMENTAL DESIGN FOR IDENTIFICATION OF DISTRIBUTED PARAMETER SYSTEMS

CHAPTER 6: OPTIMUM EXPERIMENTAL DESIGN FOR IDENTIFICATION OF DISTRIBUTED PARAMETER SYSTEMS

6.1 INTRODUCTION

In this Chapter, the problem of optimal experimental design for parameter estimation in distributed parameter systems is studied. The design variables considered are the boundary perturbation and the spatial location of measurement sensors. It is shown that suitable choice of these variables leads to improved parameter accuracy. In Section 6.2 we derive a design procedure for a general distributed parameter system. In the following two Sections we give examples to illustrate this design procedure. The first example is concerned with the optimal location of measurement sensors for estimating the velocity of propagation and the damping coefficient of a vibrating string. The second example considers the optimal design of both the boundary

perturbation and location of sensors for estimating parameters of a heat-diffusion process.

6.2 PROBLEM FORMULATION

We consider a general class of systems described by linear or non-linear hyperbolic or parabolic vector partial differential equations of the form:

$$\frac{\partial z}{\partial t} = f(t,x,z,\frac{\partial z}{\partial x}, \frac{\partial^2 z}{\partial x^2},\theta) \tag{6.2.1}$$

where f is a known function continuously differentiable with respect to x and t and twice continuously differentiable with respect to the remaining arguments.

The initial state of the system is given by:

$$z(0,x) = z_0(x) \quad \text{for } x \in \Omega \tag{6.2.2}$$

where Ω is a fixed spatial domain on which (6.2.1) holds.

The boundary conditions are assumed to be represented by:

$$g(t,z,\frac{\partial z}{\partial x},u) = 0 \quad \text{for } x \in \partial\Omega \tag{6.2.3}$$

where $\partial\Omega$ denotes the boundary of Ω and where $u(t)$, $t \in [0,T]$ is a controllable boundary perturbation.

It is further assumed that the state functions $z_i(x,t)$ are not directly measurable; instead only certain scalar functions of $z_i(x,t)$ are available at restricted measurement locations. In addition, the observations are corrupted by noise. Thus the observations can be described

by:

$$y_j(t,x_j) = h_j(t,z(t,x_j)) + e_j(t,x_j) \qquad (6.2.4)$$

for $t \in [0,T]$, where $x_j \in \Omega_s$ ($\Omega_s \in \Omega$ is the part of spatial domain where measurements can be made) and $e_j(t,x_j)$ denotes the measurement noise.

For simplicity, we assume that the measurement noise is Gaussian, spatial uncorrelated and white. The covariance is:

$$E\{e_j(t,x_j)e_k(s,x_k)\} = \sigma^2\delta_{jk}\delta(t-s) \qquad (6.2.5)$$

where δ_{jk} and $\delta(t-s)$ denote the Kronecker and Dirac delta functions, respectively.

Based on a set of N observation over [0,T], the information matrix (Goodwin & Payne [G6]) for the parameter θ is given by:

$$M = \sum_{j=1}^{N} \frac{1}{\sigma^2} \int_0^T \frac{\partial z(t,x_j)}{\partial \theta}^T \frac{\partial h_j(t,z(t,x_j))}{\partial z}^T \cdot \frac{\partial h_j(t,z(t,x_j))}{\partial z} \frac{\partial z(t,x_j)}{\partial \theta} dt \qquad (6.2.6)$$

where $\partial z(t,x)/\partial\theta$ has ij element $\partial z_i(t,x)/\partial\theta_j$ and satisfies the following set of partial differential equations:

$$\frac{\partial}{\partial t}\left[\frac{\partial z}{\partial \theta}\right] = \frac{\partial f}{\partial z}\left[\frac{\partial z}{\partial \theta}\right] + \frac{\partial f}{\partial z'}\left[\frac{\partial z'}{\partial \theta}\right] + \frac{\partial f}{\partial z''}\left[\frac{\partial z''}{\partial \theta}\right] + \frac{\partial f}{\partial \theta}$$

$$(6.2.7)$$

where $z', \partial z'/\partial\theta, z''$ and $\partial z''/\partial\theta$ denote:

$$\frac{\partial z}{\partial x}, \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial \theta}\right), \frac{\partial^2 z}{\partial x^2} \text{ and } \frac{\partial^2}{\partial x^2}\left(\frac{\partial z}{\partial \theta}\right)$$

respectively.

The initial state for (6.2.7) is given by:

$$\frac{\partial z}{\partial \theta}(0,x) = 0 \quad \text{for } x \in \Omega \qquad (6.2.8)$$

and with boundary condition:

$$\frac{\partial g}{\partial z}\left[\frac{\partial z}{\partial \theta}\right] + \frac{\partial g}{\partial z'}\left[\frac{\partial z'}{\partial \theta}\right] = 0 \quad \text{for } x \in \partial\Omega \qquad (6.2.9)$$

It is clear from Equations (6.2.1) to (6.2.9) that M is a deterministic function of the boundary perturbation $u(t)$ for $t \in [0,t]$ and of the spatial location of the measurement sensors $x_1, x_2, \ldots, x_N$.

If we use a scalar function (the determinant is normally used; see, for example, Goodwin and Payne [G6]) of the information matrix as a measure of the estimation accuracy, then it is evident that we can, in principle, choose $u(t)$ for $t \in [0,T]$ and $x_1, \ldots, x_N$ subject to constraints so that the accuracy is maximized.

In general, the non-linear partial differential equations will require numerical techniques for solution. Thus, little more can be said about the experiment design problem for distributed systems without considering a specific example. In the next two Sections, we carry through typical designs for a parabolic system and a hyperbolic system. We choose a very simple example from each system with a known analytic solution to illustrate the design procedure.

6.3 TYPICAL DESIGN FOR A HYPERBOLIC SYSTEM

There are many interesting applications for systems described by hyperbolic partial differential equations. A well-known example is the damped vibratory system. In particular, we examine the optimal sensor-location problem for a damped vibrating string. Consider a light elastic string of length π under tension and fixed at both ends. We assume that only transverse vibration takes place and the damping factor is proportional to the velocity of the string. The displacement z at any position along the string and at the time t satisfies the following one-dimensional partial differential equation:

$$\frac{\partial^2 z}{\partial t^2} + 2b\,\frac{\partial z}{\partial t} = a^2\frac{\partial^2 z}{\partial x^2} \tag{6.3.1}$$

where a is the wave velocity of propagation along the string and b is the damping coefficient. We further assume that the system satisfies the following boundary and initial conditions:

$$z(0,t) = z(\pi,t) = 0 \tag{6.3.2}$$

$$z(x,0) = f(x) \tag{6.3.3}$$

$$z_t(x,0) = g(x) \tag{6.3.4}$$

where: $$z_t(x,0) = \left.\frac{\partial z(x,t)}{\partial t}\right|_{t=0}$$

$f(x)$ and $g(x)$ are arbitrary functions of x and could be chosen freely by the experimenter. The objective is to choose x such that the information return is optimal for estimating parameters a and b. The general solution to systems described by Equations (6.3.1) to (6.3.4) is:

$$z(x,t) = \exp(-bt)\sum_{n=0}^{\infty}[f_n\cos\omega_n t+(g_n+bf_n)\frac{\sin\omega_n t}{\omega_n}]\sin nx \tag{6.3.5}$$

where: $\omega_n = [(na)^2 - b^2]^{\frac{1}{2}}$

$$g_n = \frac{1}{\pi}\int_{-\pi}^{\pi} g(x)\sin nx\,dx$$

$$f_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx$$

To simplify the design procedure, we choose $f(x) = 0$, $g(x) = \omega_k \sin kx$, where $\omega_k = [(ka)^2-b^2]^{\frac{1}{2}}$ and k is an integer constant; then (6.3.5) is reduced to:

$$\begin{aligned} z(x,t) &= \exp(-bt)\sin\omega_k t\,\sin^2 kx \\ &= \alpha\,\sin^2 kx \end{aligned} \tag{6.3.6}$$

where: $$\alpha = \exp(-bt)\sin\omega_k t \tag{6.3.7}$$

Now, we assume that measurement is made at position x_1 along the string; i.e. we observe:

$$y(x,t) = z(x_1,t) + e(x_1,t) \tag{6.3.8}$$

where $e(x_1, t)$ denotes the measurement noise (assumed Gaussian, spatial uncorrelated and white) with variance σ^2. Following Section 6.2, the average information matrix is given by:

$$\bar{M} = \frac{1}{T}\int_0^T \frac{1}{\alpha^2}\begin{bmatrix} \left(\frac{\partial z}{\partial a}\right)^2 & \left(\frac{\partial z}{\partial a}\right)\left(\frac{\partial z}{\partial b}\right) \\ \left(\frac{\partial z}{\partial a}\right)\left(\frac{\partial z}{\partial b}\right) & \left(\frac{\partial z}{\partial b}\right)^2 \end{bmatrix} dt \quad (6.3.9)$$

Substituting $z(x,t)$ from (6.3.6), $\bar{M}$ becomes:

$$\bar{M} = \sin^4 kx_1 \frac{1}{T\alpha^2}\int_0^T \begin{bmatrix} \left(\frac{\partial \alpha}{\partial a}\right)^2 & \left(\frac{\partial \alpha}{\partial a}\right)\left(\frac{\partial \alpha}{\partial b}\right) \\ \left(\frac{\partial \alpha}{\partial a}\right)\left(\frac{\partial \alpha}{\partial b}\right) & \left(\frac{\partial \alpha}{\partial b}\right)^2 \end{bmatrix} dt \quad (6.3.10)$$

To choose the optimal measurement position x_1, we maximize $\det \bar{M}$. We observe that α is independent of x_1; hence $\det \bar{M}$ is maximized by choosing:

$$x_1 = \frac{n\pi}{2k} \quad n=1,3,5\ldots, \leq k \quad (6.3.11)$$

We note that the positions given by (6.3.11) correspond to the antinodes of the damped vibrating string and that this is consistent with one's general intuition for such a system. We also remark that the choice of initial

conditions in this example is for convenience only, and in many situations they should be treated as design variables. In the next Section, we consider a heat-diffusion process in which both the boundary perturbation and location of sensors are treated as design variables.

6.4 TYPICAL DESIGN FOR A PARABOLIC SYSTEM

From among the many interesting systems described by parabolic partial differential equations, we choose a simple heat-diffusion process. Consider a semi-infinite thin radiating rod, thin enough so that radial temperature gradients are unimportant. The temperature z at any position x and at the time t in the rod satisfies the following one-dimensional partial differential equation:

$$k \frac{\partial^2 z}{\partial x^2} = \frac{\partial z}{\partial t} + \mu z \qquad (6.4.1)$$

where k is the thermal diffusivity and μ is the radiation constant.

We assume that the end of the rod is exposed to a heat source whose temperature varies sinusoidally with time. We have assumed sinusoidal variations as this leads to an analytic solution to the partial differential equation, and this form of boundary perturbation is easily implemented in practice (Sidles & Danielson [S3], Leden et al. [L1]). The boundary conditions therefore become:

$$z(0,t) = A_1 \sin\omega t \qquad (6.4.2)$$

$$z(\infty,t) = 0 \qquad (6.4.3)$$

where ω is the frequency of perturbation.

Details of practical ways of simulating these boundary conditions on a finite-length rod are given in the aforementioned two references.

The general solution to the partial differential equation is:

$$z(t,x) = A_1 \exp(-\alpha x) \sin(\omega t - \beta x) \tag{6.4.4}$$

where:

$$\alpha = \left[\frac{(\mu^2+\omega^2)^{\frac{1}{2}} + \mu}{2k}\right]^{\frac{1}{2}} \tag{6.4.5}$$

$$\beta = \left[\frac{(\mu^2+\omega^2)^{\frac{1}{2}} - \mu}{2k}\right]^{\frac{1}{2}} \tag{6.4.6}$$

Now we assume that one measurement is made at position x_1 on the rod; i.e., we observe:

$$y(t,x_1) = z(t,x_1) + e(t,x_1) \tag{6.4.7}$$

where $e(t,x_1)$ denotes the measurement noise (assumed Gaussian, spatial uncorrelated and white). The noise is assumed stationary with variance σ^2.

Following Section 6.2, the information matrix is given by:

$$M = \int_0^T \frac{1}{\sigma^2} \begin{bmatrix} \left(\frac{\partial z}{\partial k}\right)^2 & \left(\frac{\partial z}{\partial k}\right)\left(\frac{\partial z}{\partial \mu}\right) \\ \left(\frac{\partial z}{\partial k}\right)\left(\frac{\partial z}{\partial \mu}\right) & \left(\frac{\partial z}{\partial \mu}\right)^2 \end{bmatrix} dt \quad (6.4.8)$$

where:

$$\frac{\partial z}{\partial t} = -A_1 x_1 \exp(-\alpha x_1) \sin(\omega t - \beta x_1) \frac{\partial \alpha}{\partial k}$$

$$-A_1 x_1 \exp(-\alpha x_1) \cos(\omega t - \beta x_1) \frac{\partial \beta}{\partial k} \quad (6.4.9)$$

$$\frac{\partial z}{\partial \mu} = -A_1 x_1 \exp(-\alpha x_1) \sin(\omega t - \beta x_1) \frac{\partial \alpha}{\partial \mu}$$

$$-A_1 x_1 \exp(-\alpha x_1) \cos(\omega t - \beta x_1) \frac{\partial \beta}{\partial \mu} \quad (6.4.10)$$

The average information matrix over n periods of $\sin(\omega\tau)$ is given by:

$$\overline{M} = \frac{1}{2\pi n \sigma^2} \int_0^{2\pi n} \begin{bmatrix} \left(\frac{\partial z}{\partial k}\right)^2 & \left(\frac{\partial z}{\partial k}\right)\left(\frac{\partial z}{\partial \mu}\right) \\ \left(\frac{\partial z}{\partial k}\right)\left(\frac{\partial z}{\partial \mu}\right) & \left(\frac{\partial z}{\partial \mu}\right)^2 \end{bmatrix} d\tau \quad (6.4.11)$$

where $\tau = \omega t - \beta x_1$.

Using the properties that:

$$\int_0^{2\pi n} \sin^2\tau \, d\tau = n\pi \qquad (6.4.12)$$

$$\int_0^{2\pi n} \cos^2\tau \, d\tau = n\pi \qquad (6.4.13)$$

$$\int_0^{2\pi n} \sin\tau \, \cos\tau \, d\tau = 0 \qquad (6.4.14)$$

we obtain the following expression for the average information matrix:

$$\bar{M}(x_1, \omega) = \frac{1}{2\sigma^2} A_1^2 \, x_1^2 \exp(-2\alpha x_1) \cdot \begin{bmatrix} \left(\frac{\partial\alpha}{\partial k}\right)^2 + \left(\frac{\partial\beta}{\partial k}\right)^2 & \frac{\partial\alpha}{\partial k}\frac{\partial\alpha}{\partial\mu} + \frac{\partial\beta}{\partial k}\frac{\partial\beta}{\partial\mu} \\ \frac{\partial\alpha}{\partial k}\frac{\partial\alpha}{\partial\mu} + \frac{\partial\beta}{\partial k}\frac{\partial\beta}{\partial\mu} & \left(\frac{\partial\alpha}{\partial\mu}\right)^2 + \left(\frac{\partial\beta}{\partial\mu}\right)^2 \end{bmatrix} \qquad (6.4.15)$$

where:

$$\frac{\partial\alpha}{\partial k} = -\frac{1}{4}\left[\frac{(\mu^2+\omega^2)^{\frac{1}{2}} + \mu}{2k}\right]^{-\frac{1}{2}} \left[\frac{(\mu^2+\omega^2)^{\frac{1}{2}} + \mu}{k^2}\right] \qquad (6.4.16)$$

$$\frac{\partial\alpha}{\partial\mu} = \frac{1}{4}\left[\frac{(\mu^2+\omega^2)^{\frac{1}{2}} + \mu}{2k}\right]^{-\frac{1}{2}} \left[\frac{\mu(\mu^2+\omega^2) + 1}{k}\right] \qquad (6.4.17)$$

$$\frac{\partial\beta}{\partial k} = -\frac{1}{4}\left[\frac{(\mu^2+\omega^2)^{\frac{1}{2}} - \mu}{2k}\right]^{-\frac{1}{2}} \left[\frac{(\mu^2+\omega^2)^{\frac{1}{2}} - \mu}{k^2}\right] \qquad (6.4.18)$$

$$\frac{\partial\beta}{\partial\mu} = \frac{1}{4}\left[\frac{(\mu^2+\omega^2)^{\frac{1}{2}} - \mu}{2k}\right]^{-\frac{1}{2}} \left[\frac{(\mu^2+\omega^2)^{\frac{1}{2}} - 1}{k}\right] \qquad (6.4.19)$$

and:

$$\left(\frac{\partial\alpha}{\partial k}\right)^2 + \left(\frac{\partial\beta}{\partial k}\right)^2 = \frac{1}{4}\,\frac{(\mu^2+\omega^2)^{\frac{1}{2}}}{k^3} \qquad (6.4.20)$$

$$\frac{\partial\alpha}{\partial k}\,\frac{\partial\alpha}{\partial\mu} + \frac{\partial\beta}{\partial k}\,\frac{\partial\beta}{\partial\mu} = -\frac{1}{4}\,\frac{\mu(\mu^2+\omega^2)^{-\frac{1}{2}}}{k^2} \qquad (6.4.21)$$

$$\left(\frac{\partial\alpha}{\partial\mu}\right)^2 + \left(\frac{\partial\beta}{\partial\mu}\right)^2 = \frac{1}{4}\,\frac{(\mu^2+\omega^2)^{-\frac{1}{2}}}{k} \qquad (6.4.22)$$

We now assume that the amplitude A_1 of the sinusoidal variation is fixed, but that x_1, the position where the measurements are taken, and ω, the perturbation frequency, can be chosen by the experimenter so as to maximize det $\overline{M}$.

We observe that α and β are independent of x_1. Hence, for each, det $\overline{M}$ is maximized by choosing:

$$x_1 = x_1^{\dagger} = \frac{1}{\alpha} \qquad (6.4.23)$$

Substituting (6.4.23) into (6.4.15) gives:

$$\bar{M}\,(x^{\dagger}{}_{1},\omega) = \frac{1}{2\sigma^2} A_1{}^2 (\tfrac{1}{\alpha})^2 \exp(-2)$$

$$\circ \begin{bmatrix} \left(\dfrac{\partial\alpha}{\partial k}\right)^2 + \left(\dfrac{\partial\beta}{\partial k}\right)^2 & \dfrac{\partial\alpha}{\partial k}\dfrac{\partial\alpha}{\partial\mu} + \dfrac{\partial\beta}{\partial k}\dfrac{\partial\beta}{\partial\mu} \\ \dfrac{\partial\alpha}{\partial k}\dfrac{\partial\alpha}{\partial\mu} + \dfrac{\partial\beta}{\partial k}\dfrac{\partial\beta}{\partial\mu} & \left(\dfrac{\partial\alpha}{\partial\mu}\right)^2 + \left(\dfrac{\partial\beta}{\partial\mu}\right)^2 \end{bmatrix}$$

(6.4.24)

Substituting (6.4.20), (6.4.21) and (6.4.22) into (6.4.24) and solving for the determinant gives:

$$\det \bar{M}(x^{\dagger}{}_{1},\omega) = \frac{A_1{}^4 \exp(-4)\ (\frac{\omega}{\mu})^2}{16\sigma^4 k^2 \mu^2 [(1+(\frac{\omega}{\mu})^2)^{\frac{1}{2}}+1]^2\ [1+(\frac{\omega}{\mu})^2]}$$

(6.4.25)

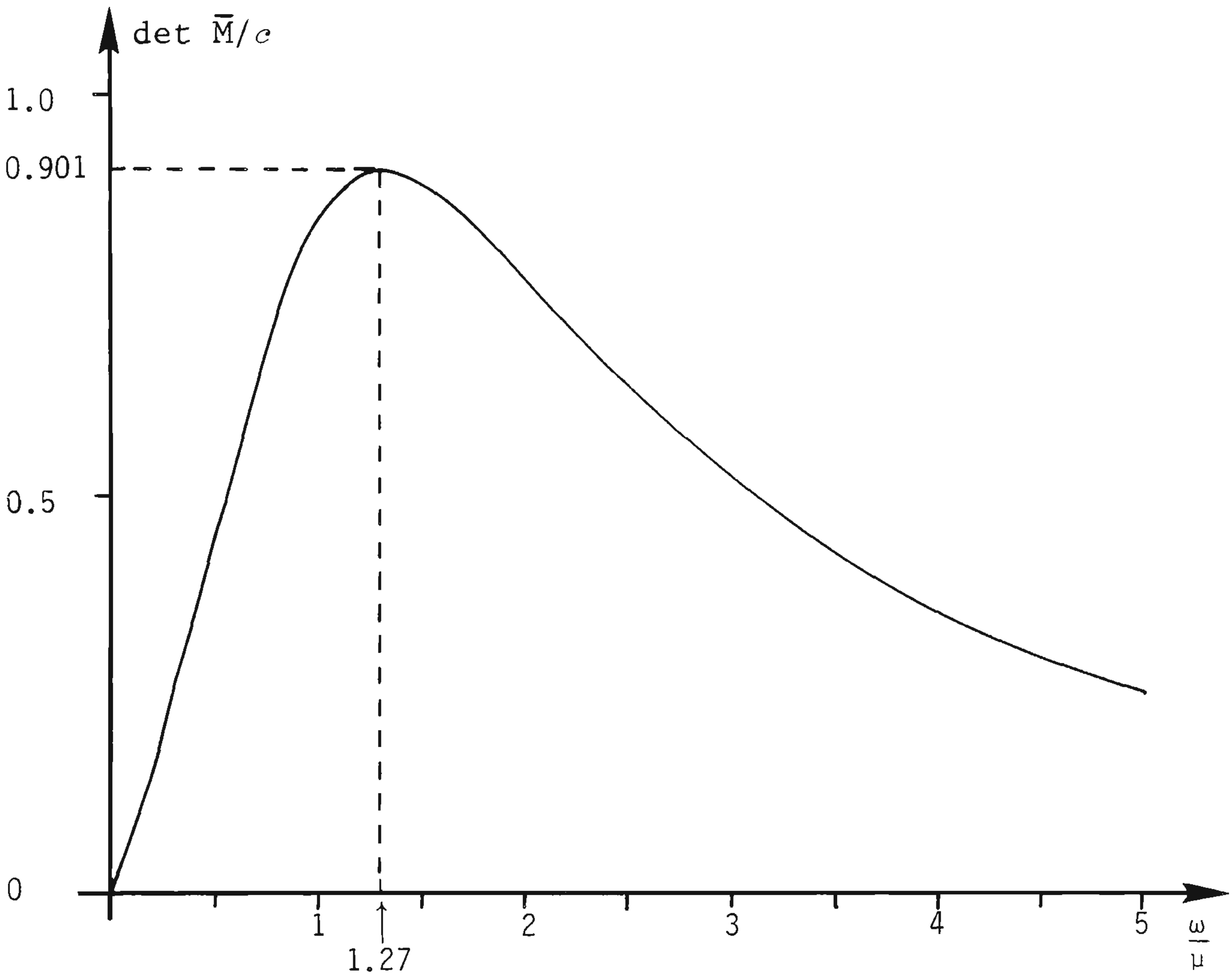
The Figure 6.1 shows the variation of det $\bar{M}(x^{\dagger}{}_{1},\omega)$ with ω.

From Equation (6.4.23) and the Figure 6.1, it can be seen that the joint optimal experimental conditions are given by:

$$x_1 = x_1{}^{\dagger\dagger} = \left[\frac{2k}{(1+1.27^2)^{\frac{1}{2}} + 1}\right]^{\frac{1}{2}} \left[\frac{1}{\mu}\right]^{\frac{1}{2}} \qquad (6.4.26)$$

$$\omega = \omega^{\dagger\dagger} = 1.27\mu \qquad (6.4.27)$$

It is evident from Figure 6.1 that the experimental conditions have a marked effect on the achievable accuracy. Thus there is considerable motivation, in practice, to choose the appropriate boundary perturbation and sensor location to optimize the information return from the experiment.



<u>Fig. 6.1</u>: Variation of det $\bar{M}$ where $c = \dfrac{A_1{}^4 \exp(-4)}{16\sigma^4 k^2 \mu^2}$

6.5 CONCLUSION

In this Chapter we presented a method to design optimal experiments for parameter estimation of a general distributed parameter system. The design procedure was illustrated with reference to a damped vibrating elastic string and a simple heat-diffusion process, and the optimal experimental conditions for these systems have also been given.

CHAPTER 7

EXTENSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

CHAPTER 7: EXTENSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

To conclude this thesis, it is appropriate to point to some remaining problems. Suggestions for extending certain results in this thesis, and areas where further research may be done, are discussed briefly below:

(1) In Chapter 3, the optimal input design was obtained by solving a set of non-linear equations (3.4.1), (3.4.2a). In Section 3.5, the non-linear equations for the first-order AR model were simple to solve, whereas the non-linear equations for the second-order AR model were solved by arbitrarily choosing the value of ω_1, and then solving for the other design variables. The non-linear equations become more complicated to solve as the order of the system increases. Iterative methods to solve non-linear equations can be found in Ortega and Rheinbold [01][02].

To overcome the problem of solving non-linear equations (3.4.1), (3.4.2a), an approach was considered where the solution to (3.4.1), (3.4.2a) could be obtained by solving two sets of linear equations. This approach posed some difficulties, and the main steps are presented below.

First note that (3.4.1) and (3.4.2a) are the real part of the following set of equations:

$$\sum_{i=1}^{\ell} g_i \, e^{jk\omega_i} = h \, e^{j\alpha_k}, \; k=0,1,\ldots,n \tag{1}$$

where:

$$g_i = \frac{m_i}{f(\omega_i)}$$

$$h = W - C_0$$

$$h_k = h \cos\alpha_k = -C_k \;, \; k=1,\ldots,n$$

For n odd, we have 2ℓ simultaneous equations and for the case n even, we introduce the additional equation:

$$\sum_{i=1}^{\ell} g_i \, e^{j(n+1)\omega} = h \, e^{j\alpha_{n+1}} \tag{2}$$

with $h \cos\alpha_{n+1} = -C_{n+1}$

Introduce the polynomial

$$P(x) = p_0 + p_1 x + p_2 x^2 + \ldots, + p_\ell x^{\ell-1} + x^\ell = 0 \tag{3}$$

where $x = e^{j\omega}$, and let x_i, i=1,..., ℓ be the roots of $P(x)$.

Now multiply the first equation of (1) by p_0, the second equation by $p_1,\ldots,$ and the $(\ell+1)$st equation by 1, and adding yields:

$$p_0\, e^{j\alpha_0} + p_1\, e^{j\alpha_1} + \ldots + p_{\ell-1}\, e^{j\alpha_{\ell-1}} = -e^{j\alpha_\ell}$$

Carrying out similarly as above, by taking the next $(\ell+1)$ equations of (1), i.e. for k=1,...,ℓ+1, and next repeating with the next $(\ell+1)$ equations of (1), we finally arrive at the following linear equations:

$$\begin{bmatrix} e^{j\alpha_0} & e^{j\alpha_1} & \cdots & e^{j\alpha_{\ell-1}} \\ e^{j\alpha_1} & & & e^{j\alpha_\ell} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ e^{j\alpha_{\ell-1}} & & \cdots & e^{j\alpha_{2\ell-2}} \end{bmatrix} \begin{bmatrix} p_0 \\ p_1 \\ \\ \\ \\ p_{\ell-1} \end{bmatrix} = - \begin{bmatrix} e^{j\alpha_\ell} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ e^{j\alpha_{2\ell-1}} \end{bmatrix}$$

the values of $p_0,\ldots,\ p_{\ell-1}$ can be obtained from this set of equations. The roots of $P(x)$ can now be determined from (3).

Next, g_i, i=1,..., ℓ can be obtained from the first ℓ equations of the set of equations (3.4.1), (3.4.2a), i.e. by solving the following linear equations:

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \cos\omega_1 & \cos\omega_2 & & \cos\omega_\ell \\ \cos 2\omega_1 & & & \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ \cos(\ell-1)\omega_1 & & \cdots & \cos(\ell-1)\omega_\ell \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ g_\ell \end{bmatrix} = \begin{bmatrix} W - C_0 \\ -C_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ -C_n \end{bmatrix}$$

With g_i, $i=1,\ldots,\ell$ known, m_i can easily be calculated.

Some of the problems encountered in the above approach were that the roots of $P(x)$ may not lie on the unit circle, and the constraints $\Sigma g_i \sin k\omega_i = h \sin\alpha_k$ are not satisfied.

Further work is needed to investigate the conditions for the roots of $P(x)$ to lie on the unit circle, and to look into the other problems mentioned in this approach.

(2) D_s-optimal and D-optimal designs have been compared in Chapter 4 (Section 4.4) for the simple case, s=1, p=2. The criterion used for comparing the two designs was the variance of the parameter estimate. Further work is required to compare the two designs when $2 \leq s \leq p$; in particular, we need to find a suitable criterion for comparing the two designs. Can we, in general, say that D_s-optimal design is always better than D-optimal design?

(3) In Chapter 4, D_s-optimal designs were considered for the case when $\bar{M}$ is non-singular. For singular $\bar{M}$, transformations used by Kiefer [K4] and Atwood [A4] in linear regression problems cannot be extended for dynamic systems. Unimodular matrices also do not result in linear transformations. Further work is required to tackle the D_s-optimal design for the case of singular $\bar{M}$.

(4) Algorithms converging to D-optimal designs have been considered by a number of authors; see, for example, Mehra [M4], Zarrop [Z3].

An iterative design algorithm which is an extension form from Atwood [A5] is presented, which holds for $p > 2$.

Following the definitions in Chapter 4 and Zarrop [Z3], we have:

$$\bar{M}\,(\xi_{k+1}) \;=\; (1-\alpha_k)\,\bar{M}\,(\xi_k) + \alpha_k\, Re\,\{h(e^{j\omega_k})\, h^*(e^{j\omega_k})$$

$\bar{M}$ is a non-singular p x p matrix.

$$\bar{M}\,(\xi_{k+1}) \;=\; (1-\alpha_k)\,\bar{M}\,(\xi_k) + \tfrac{1}{2}\alpha_k\, h(e^{j\omega_k})\, h^*(e^{j\omega_k}) + \tfrac{1}{2}\alpha_k\, \bar{h}(e^{j\omega_k})\, h^T(e^{j\omega_k})$$

$$\frac{\det \bar{M}\,(\xi_{k+1})}{\det \bar{M}\,(\xi_k)} = (1+\beta)^{-p}\,\{1 + \beta d(\omega_k,\xi_k) + \beta^2 g(\omega_k,\xi_k)\}$$

$$d(\omega,\xi) = h^*(e^{j\omega})\,\bar{M}^{-1}(\xi)\,h(e^{j\omega})$$

$$g(\omega,\xi) = \tfrac{1}{4}\{d^2(\omega,\xi) - |d_1(\omega,\xi)|^2\}$$

$$d_1(\omega,\xi) = h^T(e^{j\omega})\ \bar{M}^{-1}(\xi)\ h(e^{j\omega})$$

and $\quad \beta = \dfrac{\alpha}{1-\alpha}$

$$\log\left[\frac{\det\ \bar{M}\ (\xi_{k+1})}{\det\ \bar{M}\ (\xi_k)}\right] = -p\ \log(1+\beta) + \log\ (1+\beta d+\beta^2 g)$$

log is defined for:

$\beta > -1$ and

$$\beta > \min\left[-\frac{2}{d+|d_1|}\ ,\ -\frac{2}{d-|d_1|}\right]$$

differentiating log $\{\det M\ (\xi_{k+1})/\det M\ (\xi_k)\}$ with respect to β and equating equal to zero gives:

$$\beta^2[g(p-2)] + \beta[d(p-1)-2g] + p - d = 0$$

$$\beta = -\frac{[d(p-1) - 2g]}{2g(p-2)} \pm A$$

β has real roots if $d(\xi_k,\xi_k) > p$

where: $\quad A = \dfrac{\sqrt{[d(p-1) - 2g]^2 - 4g(p-2)(p-d)}}{2g(p-2)}$

Now, $A > 0$ and $A > \dfrac{d(p-1) - 2g}{2g(p-2)}$

the positive root of β maximizes the function and the negative root minimizes the function.

If β is chosen as shown above, then det $\bar{M}$ (ξ_k) converges to the optimum.

det $\bar{M}$ (ξ_k), $k \geq k_0$ is a monotone-increasing sequence and is bounded above, so it converges, i.e.:

$$\lim_{k\to\infty} \det \bar{M}(\xi_k) = \det \bar{M}(\xi') \leq \det \bar{M}(\xi^{\dagger})$$

To show that det $\bar{M}$ (ξ_k) converges to the optimum, assume the contrary, i.e. det $\bar{M}$ $(\xi') \neq$ det $\bar{M}$ $(\xi^{\dagger})$; then it follows from the equivalence theorem (Mehra [M4]), that $d(\xi_k, \xi_k) - p \geq \eta > 0$.

Because the sequence converges, there exists a small positive number γ, such that, for $k > k'$ det $\bar{M}$ (ξ_{k+1}) - det $\bar{M}$ $(\xi_k) \leq \gamma$. For any $\Delta > 0$, there is a γ such that $d(\omega_k, \xi_k) - p = \delta_k \leq \Delta$.

In this way we can specify k' such that d-p will be less than any Δ. So if we choose $\Delta < \eta$ there is a contradiction. Therefore det $\bar{M}$ (ξ_k) converges to the optimal design.

Further work is required to consider the case when p=1,2.

(5) The expression for det $\overline{M}$ becomes more complicated as the order of the system increases. Zarrop [Z3] has shown that a considerable simplification occurs in deriving the expression for det $\overline{M}$ if either the parameters in $A(z^{-1})$ or $B(z^{-1})$ are known. Using these expressions for det $\overline{M}$, we present a conjecture for a plausible sequential design algorithm:

(i) estimate the numerator parameters assuming the denominator parameters are known (one can start with the nominal values of parameters);

(ii) estimate the denominator parameters using the numerator parameter values from (i);

(iii) update the design.

Keep switching between the numerator and denominator until a D-optimal design is achieved.

The above algorithm requires further work. For example, it is required at least to show that the algorithm converges to a D-optimal design.

(6) In Chapter 5, an expression for the lower bound of the efficiency of minimal uniform designs is given by (5.3.3a), which is:

$$\eta^p \geq \frac{p^{-p}\ |G(p)|\,|G^*(p)|}{\sum_i \beta_i \ \max_j \{|G_j|\,|G_j^*|\}} \tag{4}$$

Since $\phi(\beta) = \sum_i \beta_i$ is a symmetric multilinear function, satisfying $0 \leq \beta_i \leq 1$, it follows from Keilson [K7] that maximum $\phi(\beta)$ occurs at points:

$$\beta = [(2\ell)^{-p},\ldots,(2\ell)^{-p}] \text{ or } [0, (2\ell-1)^{-p},\ldots,(2\ell-1)^{-p}]$$

$$\text{or } [0,0,(2\ell-2)^{-p},\ldots,(2\ell-2)^{-p}] \text{ etc.}$$

Hence:

$$\eta^p = \frac{p^{-p}}{\max \binom{v}{p} v^{-p}} \quad \frac{|G(p)|\,|G^*(p)|}{\max_i\{|G_i|\,|G_i^*|\}} \tag{5}$$

where: $p \leq v \leq 2\ell$

Now we have the ratio:

$$\frac{\binom{v}{p}\ v^{-p}}{\binom{v-1}{p}\ (v-1)^{-p}} = \frac{(1 - 1/v)^p}{(1 - p/v)} \geq 1$$

The latter inequality is obtained by expanding $(1 - 1/v)^p$ Therefore in (5) the maximum occurs when v is as large as possible. Hence:

$$\eta^p \geq \frac{p^{-p}}{\binom{2\ell}{p}(2\ell)^{-p}} \quad \frac{|G(p)|\,|G^*(p)|}{\max_i\{|G_i|\,|G_i^*|\}}$$

It has already been mentioned in Chapter 5 that:

$$|G(p)|\,|G^*(p)| \geq \max_i \{|G_i|\,|G_i^*|\}$$

when G_i is of the form of (5.3.2b). This ratio between $|G(p)|\,|G^*(p)|$ and $\max\limits_i \{|G_i|\,|G_i^*|\}$ is not known when G_i is not of the form given by (5.3.2b).

It would be of interest to find out conditions under which G_i is of the form given by (5.3.2b). In this case, the lower bound will only depend on the number of parameters and η^p is given by:

$$\frac{p^{-p}}{\binom{2\ell}{p}(2\ell)^{-p}}$$

Another extension is to compare uniform designs with different numbers of input frequencies, and to obtain information on the maximum possible improvement in efficiency when the number of input frequencies is increased in a uniform design. For specific examples, the procedure required is identical to that presented in Chapter 5. However, similar problems arise in the general case.

(7) Most of the problems on input design have considered the general model given by (2.2.1). The ARMA model given by:

$$y_k = \frac{B(z^{-1})}{A(z^{-1})} u_{k-d} + \frac{1}{A(z^{-1})} e_k$$

is a special case of the general model where the system t.f and the noise t.f have some common parameters. The input design problem for the ARMA model has not, as yet, been considered. It would be of interest to investigate the optimal input design problem for this model, under both the input and output power constraints.

(8) Contrary to lumped-parameter systems, in which the problems of optimal experimental design are well developed, similar investigations for DPS seem to be, as yet, at an early stage of development. More recently Rafajlowicz [R3] has considered the experimental design problem for a class of DPS, described by a linear, parabolic partial differential equation, to estimate the system's eigenvalues rather than its parameters. He gave conditions for optimality of both input signal spectral density and a probability measure corresponding to positions of sensors.

Further investigations are desirable to determine efficient experimental design procedures for different classes of DPS with not only measurement noise, but also process noise.

(9) This thesis has dealt mainly with the problem of input signal design for parameter estimation. However, some other aspects of experimental design, for example, sampling rate design, choice of presampling filter, have not been considered in this thesis. A number of authors have studies these aspects; see, for example, Payne [P1], Payne et al. [P5], and Ng and Goodwin [N8]. Goodwin et al. [G10] have shown that to achieve maximal return from an experiment, coupled design of all the experimental conditions, namely the test signal, sampling intervals and filters, should be carried out simultaneously. This is another area of future research.

AUTHOR'S PUBLICATIONS

AUTHOR'S PUBLICATIONS

Qureshi, Z.H., Ng, T.S. and Goodwin, G.C., "Optimum experimental design for identification of distributed parameter systems", *International Journal of Control*, 31, 1, 21-29, January 1980.

Ng. T.S. and Qureshi, Z.H., "Optimal experiment design for autoregressive model with output power constraints", *IEEE Transactions on Automatic Control*, AC-26, 3, 739-742, June 1981.

Ng, T.S. and Qureshi, Z.H., "Optimal input design for parameter estimation for an autoregressive model with output power constraints", *IFAC Congress*, August 24-28, 1981, Kyoto Japan.

Qureshi, Z.H. and Ng, T.S., "Optimal input design for dynamic system parameter estimation - the D_s-optimality case", to be published in *SIAM Journal on Control and Optimization*, September 1982.

REFERENCES

REFERENCES

[A1] Aoki, M. and Staley, R.M., "On input signal synthesis in parameter identification", *Automatica,* 6, 431-440, 1970.

[A2] Aström, K.J. and Bohlin, T., "Numerical identification of linear dynamic systems from normal operating records", *IFAC Symposium,* Teddington, England, September 1965.

[A3] Aström, K.J. and Eykhoff, P., "System identification - A survey", *Automatica,* 7, 123-162, 1971.

[A4] Atwood, C.L., "Optimal and efficient designs of experiments", *Ann. Math. Statist.,* 40, 5, 1570-1602, 1969.

[A5] Atwood, C.L., "Sequences converging to D-optimal designs of experiments", *Ann. Statist.* 1, 2, 342-352, 1973.

[A6] Aström, K.J., Jury, E.I. and Agniel, R.G., "A numerical method for the evaluation of complex integrals", *IEEE Trans. Automatic Control,* AC-15, 468-471, 1970.

[B1] Berger, C.S., "Synthesis of input signals for parameter identification using moving average filter", *Proc. IEE,* 126, 12, 1311-1315, 1979.

[C1] Carpenter, W.T., Wozny, M.J. and Goodson, R.E., "Distributed parameter identification using the method of characteristics", ASME Trans., *J. Dyn. Syst. Meas. and Control,* 93-G, 2, 73-78, 1971.

[E1] Eykhoff, P., *System Identification.* John Wiley, London, 1974.

[F1] Federov, V.V., *Theory of Optimal Experiments,* Academic Press, New York, 1972.

[G1] Goodson, R.E. and Klein, R.E., "A definition and some results for distributed parameter system observability", *IEEE Trans. Automatic Control,* AC-15, 165-174, 1970.

[G2] Goodwin, G.C., "Input synthesis for minimum covariance state and parameter estimation", *Electronics Letters,* 5, 21, 16 October, 1969.

[G3] Goodwin, G.C., "Optimal input signals for nonlinear system identification", *Proc. IEE,* 118, 7, 922-926, 1971.

[G4] Goodwin, G.C., Murdoch, J.C. and Payne, R.L., "Optimal test signal design for linear S.I.S.O. system identification", *Int. J. Control,* 17, 1, 45-55, 1973.

[G 5] Goodwin, G.C. and Payne, R.L., "Design and characterisation of optimal test signals for linear single input-single output parameter estimation", Paper TT-1, *Proc. 3rd IFAC Symposium on Identification and System Parameter Estimation*, The Hague, Delft, June 1973.

[G 6] Goodwin, G.C. and Payne, R.L., *Dynamic System Identification: Experiment design and data analysis.* Academic Press, New York, 1977.

[G 7] Graupe, D., *Identification of Systems.* Van Nostrand Reinhold, New York, 1972.

[G 8] Gustavsson, I., "Survey of applications of identification in chemical and physical processes", Paper S-3, *Proc. 3rd IFAC Symposium on Identification and System Parameter Estimation*, The Hague, Delft, June 1973.

[G 9] Gustavsson, I., Ljung, L. and Söderström, T., "Identification of processes in closed loop - Identifiability and accuracy aspects", *Automatica*, 13, 59-75, 1977.

[G10] Goodwin, G.C., Zarrop, M.B. and Payne, R.L., "Coupled design of test signals, sampling intervals, and filters for system identification", *IEEE Trans. Automatic Control*, AC-19, 6, 748-752, 1974.

[J 1] Jury, E.I., *Theory and Application of the Z-Transform Method.* John Wiley, New York, 1964.

[K 1] Kiefer, J., "Optimal experimental design", *J. Royal Statist. Soc.*, Ser.B, 21, 272-319, 1959.

[K 2] Kiefer, J. and Wolfowitz, J., "Optimum designs in regression problems", *Ann. Math. Statist.*, 30, 271-294, 1959.

[K 3] Kiefer, J. and Wolfowitz, J., "The equivalence of two extremum problems", *Canadian J. Math.*, 12, 363-366, 1960.

[K 4] Kiefer, J., "Optimum designs in regression problems, II", *Ann. Math. Statist.*, 32, 298-325, 1961.

[K 5] Keviczky, L. and Banyasz, Cs., "On input signal synthesis for linear discrete-time systems", Paper TT-2, *Proc. 3rd IFAC Symposium on Identification and System Parameter Estimation*, The Hague, Delft, June 1973.

[K 6] Kubrusly, C.S., "Distributed parameter system identification: A survey", *Int. J. Control*, 26, 4, 509-535, 1977.

[K 7] Keilson, J., "A theorem on optimum allocation for a class of symmetric multilinear return functions", *J. Math. Appl.*, 15, 269-272, 1969.

[L1] Leden, B., Hamzer, M.H. and Sheirah, M.A., "Diffusion methods for estimation of thermal diffusivity of a heat diffusion process", Paper TD-1, *Proc. 3rd IFAC Symposium on Identification and System Parameter Estimation*, The Hague, Delft, June 1973.

[L2] Levadi, V.S., "Design of input signals for parameter estimation", *IEEE Trans. Automatic Control*, AC-11, 2, 205-211, 1966.

[L3] Levin, M.J., "Optimum estimation of impulse response in the presence of noise", *IRE Trans. Circuit Theory*, CT-7, 50-56, 1960.

[L4] Litman, S. and Huggins, W.H., "Growing exponentials as a probing signal for system identification", *Proc. IEEE*, 31, 6, 917-923, 1963.

[M1] Mehra, R.K., "Optimal inputs for linear system identification", *IEEE Trans. Automatic Control*, AC-19, 3, 192-200, 1974.

[M2] Mehra, R.K., "Frequency domain synthesis of optimal inputs for linear system parameter estimation", Div. Eng. and Appl. Physics, Harvard University, Tech. Report 645, July 1973.

[M3] Mehra, R.K., "Synthesis of optimal inputs for MIMO systems with process noise", Div. Eng. and Appl. Physics, Harvard University, Tech. Report 649, February 1974.

[M4] Mehra, R.K., "Optimal input signals for parameter estimation in dynamic systems - Survey and new results", *IEEE Trans. Automatic Control*, AC-19, 6, 753-768, 1974.

[M5] Mehra, R.K. and Lainiotis, D.G., *System Identification: Advances and case studies.* Academic Press, New York, 1976.

[M6] Melsa, J.L. and Sage, A.P., *An Introduction to Probability and Stochastic Process.* Prentice Hall, Englewood Cliffs, New Jersey, 1973.

[N1] Nahi, N.E., *Estimation Theory and Applications.* John Wiley, New York, 1969.

[N2] Nahi, N.E. and Napjus, G.A., "Design of optimal probing signals for vector parameter estimation", Paper W9-5, *Preprints IEEE Conference on Decision and Control*, Miama, Florida, 1971.

[N3] Nahi, N.E. and Wallis, D.E.Jr., "Optimal inputs for parameter estimation in dynamic systems with white observation noise", Paper IV-A5, *Proc. JACC*, Boulder, Colorado, 1969.

[N4] Ng, T.S., Goodwin, G.C. and Payne, R.L., "On maximal accuracy estimation with output power constraints", *IEEE Trans. Automatic Control,* AC-22, 1, 133-134, 1977.

[N5] Ng, T.S., Goodwin, G.C. and Söderström, T., "Optimal experiment design for linear systems with input-output constraints", *Automatica,* 13, 571-577, 1977.

[N6] Ng. T.S. and Qureshi, Z.H., "Optimal experiment design for autoregressive model with output power constraints", *IEEE Trans. Automatic Control,* AC-26, 3, 739-742, 1981.

[N7] Nieman, R.E., Fisher, D.G. and Seborg, D.E., "A review of process identification and parameter estimation techniques", *Int. J. Control,* 13, 2, 209-264, 1971.

[N8] Ng, T.S. and Goodwin, G.C., "On optimal choice of sampling strategies for linear system identification", *Int. J. Control,* 23, 4, 459-475, 1976.

[O1] Ortega, J.M. and Rheinboldt, W.C., *Iterative Solution of Nonlinear Equations in Several Variables.* Academic Press, New York, 1970.

[O2] Ortega, J.M., *Numerical Analysis.* Academic Press, New York, 1972.

[P1] Payne, R.L., *Optimal Experiment Design for Dynamic System Identification,* Ph.D. Thesis, Imperial College, London, 1974.

[P2] Payne, R.L. and Goodwin, G.C., "Simplification of frequency domain experiment design for linear single input single output systems", Pub.74/3, Dept. of Computing and Control, Imperial College, London, 1974.

[P3] Perdreauville, F.J. and Goodson, R.E., "Identification of systems described by partial differential equations", ASME Trans., *J. Basic Eng.,* 88-D, 2, 463-468, 1966.

[P4] Polis, M.P., Goodson, R.E. and Wozny, M.J., "On parameter identification for distributed parameter systems using Galerkin's Criterion", *Automatica,* 9, 53-64, 1973.

[P5] Payne, R.L., Goodwin, G.C. and Zarrop, M.B., "Frequency domain approach for designing sampling rates for system identification", *Automatica,* 11, 2, 189-193, 1975.

[Q1] Qureshi, Z.H., Ng. T.S. and Goodwin, G.C., "Optimum experimental design for identification of distributed parameter systems", *Int. J. Control,* 31, 1, 21-29, 1980.

[Q2] Qureshi, Z.H. and Ng, T.S., "Optimal input design for dynamic system parameter estimation - The D_s-optimality case", to appear in *Siam J. Control and Optimization,* September 1982.

[R1] Rault, A., "Identification applications to aeronautics", Paper S-2, *Proc. 3rd IFAC Symposium on Identification and System Parameter Estimation,* The Hague, Delft, June 1973.

[R2] Rothenberg, T.J., "Identification in parametric models", *Econometrica,* 39, 3, 577-591, 1971.

[R3] Rafajlowicz, E., "Design of experiments for eigenvalue identification in distributed-parameter systems", *Int. J. Control,* 34, 6, 1079-1094, 1981.

[S1] Sage, A.P. and Melsa, J.L., *System Identification.* Academic Press, New York, 1971.

[S2] Seinfeld, J.H., "Identification of parameters in partial differential equations", *Chemical Engineering Science,* 24, 1, 65-74, 1969.

[S3] Sidles, P.H. and Danielson, G.C., "Thermal diffusivity of metals at high temperatures", *J. Applied Physics,* 25, 1, 58-66, 1954.

[S4] Söderström, T., Ljung, L. and Gustavsson, I., "On the accuracy of identification and the design of identification experiments", Report 7428, Dept. of Automatic Control, Lund Institute of Technology, Sweden, December 1974.

[S5] Söderström, T., Ljung, L. and Gustavsson, I., "On the accuracy problem in identification", *Proc. 6th IFAC Congress,* Boston, Mass., 1975.

[S6] Söderström, T., "Notes on the design of optimal identification experiments", Report UPTEC 7563R, Institute of Technology, Uppsala University, Sweden, September 1975.

[S7] Stoica, P. and Söderström, T., "A useful input parameterization for optimal experiment design", Internal Report, Dept. of Automatic Control and Systems Analysis, Institute of Technology, Uppsala University, Sweden, 1981.

[S8] St. John, R.C. and Draper, N.R., "D-optimality for regression designs: A review", *Technometrics,* 17, 1, 15-23, 1975.

[S9] Saff, E.B. and Snider, A.D., *Fundamentals of Complex Analysis for Mathematics, Science and Engineering.* Prentice Hall, New Jersey, 1976.

[T1] Tse, E., "Information matrix and local identifiability of parameters", Paper 20-3, *Proc. JACC,* Columbus, Ohio, 1973.

[T2] Tzafestas, S.G. and Nightingale, J.M., "Concerning optimal filtering theory of linear distributed parameter systems", *Proc. IEE,* 115, 11, 1737-1742, 1968.

[U1] Upadhyaya, B.R. and Sorenson, H.W., "Synthesis of linear stochastic signals in identification problems", *Automatica,* 13, 615-622, 1977.

[V1] Van den Bos, A., "Selection of periodic test signals for estimation of linear dynamic systems", Paper TT-3, *Proc. 3rd IFAC Symposium on Identification and System Parameter Estimation,* The Hague, Delft, June 1973.

[V2] Viort, B., "D-optimal designs for dynamic models: Part 1, Theory; Part 2, Application", Tech. Reports 314, 316, Dept. of Statistics, University of Wisconsin, 1972.

[W1] Wynn, H.P., "The sequential generation of D-optimum experimental designs", *Ann. Math. Statist.,* 41, 5, 1655-1664, 1970.

[W2] Wynn, H.P., "Results in the theory and construction of D-optimum experimental designs", *J. Royal Statist. Soc.,* Ser.B, 34, 133-147, 1972.

[Y1] Yu, T.K. and Seinfeld, J.H., "Observability and optimal measurement location in linear distributed parameter systems", *Int. J. Control,* 18, 4, 785-799, 1973.

[Z1] Zarrop, M.B. and Goodwin, G.C., "Comments on optimal inputs for linear system identification", *IEEE Trans. Automatic Control,* AC-20, 2, 299-300, 1975.

[Z2] Zarrop, M.B., "A Chebyshev system approach to optimal input design", *IEEE Trans. Automatic Control,* AC-24, 5, 687-698, 1979.

[Z3] Zarrop, M.B., *Optimal Experiment Design for Dynamic System Identification.* Springer-Verlag, Berlin, 1979.