SYSTEM IDENTIFICATION:

Theory for the User

Lennart Ljung

University of Linköping
Sweden

P T R Prentice Hall, Englewood Cliffs, New Jersey 07632

Library of Congress Cataloging-in-Publication Data

Ljung, Lennart. (date)
System identification

Bibliography: p.
Includes index.
1. System identification. I. Title.
QA402.L59 1987 003 86-22712
ISBN 0-13-881640-9

Editorial/production supervision and
interior design: Gloria Jordan
Cover design: Ben Santora
Manufacturing buyer: S. Gordon Osbourne

Prentice-Hall, Inc.
A Division of Simon & Schuster
Englewood Cliffs, New Jersey 07632

Printed in the United States of America

10 9 8 7

ISBN 0-13-881640-9

Prentice-Hall International (UK) Limited, *London*
Prentice-Hall of Australia Pty. Limited, *Sydney*
Prentice-Hall Canada Inc., *Toronto*
Prentice-Hall Hispanoamericana, S.A., *Mexico*
Prentice-Hall of India Private Limited, *New Delhi*
Prentice-Hall of Japan, Inc., *Tokyo*
Prentice-Hall of Southeast Asia Pte. Ltd., *Singapore*
Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*

CONTENTS

part ii: methods

part iii: user's choices

PREFACE

System identification is a diverse field that can be presented in many different ways. The subtitle, *Theory for the User*, reflects the attitude of the present treatment. Yes, the book is about theory, but the focus is on theory that has direct consequences for the understanding and practical use of available techniques. My goal has been to give the reader a firm grip on basic principles so that he or she can confidently approach a practical problem, as well as the rich and sometimes confusing literature on the subject.

Stressing the utilitarian aspect of theory should not, I believe, be taken as an excuse for sloppy mathematics. Therefore, I have tried to develop the theory without cheating. The more technical parts have, however, been placed in appendixes or in asterisk-marked sections, so that the reluctant reader does not have to stumble through them. In fact, it is a redeeming feature of life that we are able to use many things without understanding every detail of them. This is true also of the theory of system identification. The practitioner who is looking for some quick advice should thus be able to proceed rapidly to Part III (User's Choices) by hopping through the summary sections of the earlier chapters.

The core material of the book should be suitable for a graduate-level course in system identification. As a prerequisite for such a course, it is natural, although not absolutely necessary, to require that the student should be somewhat familiar with dynamical systems and stochastic signals. The manuscript has been used as a text for system identification courses at Stanford University, the Massachusetts Institute of Technology, Yale University, the Australian National University and the Univer-

sities of Lund and Linköping. Course outlines, as well as a solutions manual for the problems, are available from the publisher.

For a course on system identification, the role of computer-based exercises should be stressed. Simulation sessions demonstrating how hidden properties of data are readily recovered by the techniques discussed in the book enhance the understanding and motivation of the material. In the problems labeled S, in Chapters 2 through 16, a basic interactive software package is outlined that should be possible to implement rather painlessly in a high-level environment. A PC-MATLAB version of this package is commercially available (see Ljung, 1986b). With such a package all basic techniques of this book can be illustrated and tested on real and simulated data.

The existing literature on system identification is indeed extensive and virtually impossible to cover in a bibliography. In this book I have tried to concentrate on recent and easily available references that I think are suitable for further study, as well as on some earlier works that reflect the roots of various techniques and results. Clearly, many other relevant references have been omitted.

Finally, some words about the structure of this book: The dependence among the different chapters is illustrated in Figure 1.13, which shows that some chapters are not necessary prerequisites for the following ones. Also, some portions contain material that is directed more toward the serious student of identification theory than to the user. These portions are put either in appendixes or in sections and subsections marked with an asterisk (*). While occasional references to this material may be encountered, it is safe to regard it as optional reading; the continuity will not be impaired if it is skipped.

The problem sections for each chapter have been organized into six groups of different problem types:

- *G problems:* These could be of General interest and it may be worthwhile to browse through them, even without intending to solve them.
- *E problems:* These are regular pencil-and-paper Exercises to check the basic techniques of the chapter.
- *T problems:* These are Theoretically oriented problems and typically more difficult than the E problems.
- *D problems:* In these problems the reader is asked to fill in technical Details that were glossed over in the text (a way to dump straightforward technicalities from the book into the solutions manual!).
- *S problems:* These develop the basic identification Software package mentioned earlier.
- *C problems:* These require a Computer. Clearly, with the software package at hand, the C problems can be complemented with a myriad of problems experimenting with identification methods and data. Such problems are not specifically listed, but the reader is encouraged to apply those techniques in an exploratory fashion.

ACKNOWLEDGMENTS

Any author of a technical book is indebted to the people who taught him the subject and to the people who made the writing possible. My interest in system identification goes back to my years as a graduate student at the Automatic Control Department in Lund. Professor Karl Johan Åström introduced me to the subject, and his serious attitude to research has always been a reference model for me. Since then I have worked with many other people who added to my knowledge of the subject. I thank, therefore, my previous coauthors (in alphabetical order) Anders Ahlén, Peter Caines, David Falconer, Farhat Fnaiech, Ben Friedlander, Michel Gevers, Keith Glover, Ivar Gustavsson, Tom Kailath, Stefan Ljung, Martin Morf, Ton van Overbeek, Jorma Rissanen, Torsten Söderström, Göte Solbrand, Eva Trulsson, Bo Wahlberg, Don Wiberg, and Zhen-Dong Yuan.

The book has developed from numerous seminars and several short courses that I have given on the subject world-wide. Comments from the seminar participants have been instrumental in my search for a suitable structure and framework for presenting the topic.

Several persons have read and used the manuscript in its various versions and given me new insights. First, I would like to mention: Michel Gevers, who taught from an early version and gave me invaluable help in revising the text; Robert Kosut and Arye Nehorai, who taught from the manuscript at Stanford and Yale, respectively; and Jan Holst, who lead a discussion group with it at Denmark's Technical University, and also gathered helpful remarks. I co-taught the course at MIT with Fred Schweppe, and his lectures as well as his comments, led to many clarifying

changes in the manuscript. Students in various courses also provided many useful comments. I mention in particular George Hart, Juan Lavalle, Ivan Mareels, Brett Ridgely, and Bo Wahlberg. Several colleagues were also kind enough to critique the manuscript. I am especially grateful to Hiro Akaike, Chris Byrnes, Peter Falb, Meir Feder, Gene Franklin, Claes Källström, David Ruppert, Torsten Söderström, Petre Stoica, and Peter Whittle.

Svante Gunnarsson and Stan Granath made the experiments described in Section 17.2, Bo Wahlberg contributed to the frequency-domain interpretations, and Alf Isaksson prepared Figure 14.4.

The preparation of the manuscript's many versions was impeccably coordinated and, to a large extent, also carried out by Ingegerd Stenlund. She had useful help from Ulla Salaneck and Karin Lönn. Marianne Anse-Lundberg expertly prepared all the illustrations. I deeply appreciate all their efforts.

Writing a book takes time, and I probably would not have been able to finish this one had I not had the privilege of sabbatical semesters. The first outline of this book was written during a sabbatical leave at Stanford University in 1980–1981. I wrote a first version of what turned out to be the last edition during a mini-sabbatical visit to the Australian National University in Canberra in 1984. The writing was completed during 1985–1986, the year I spent at MIT. I thank Tom Kailath, Brian Anderson and Sanjoy Mitter (and the U.S. Army Research Office under contract DAAG-29-84-K-005) for making these visits possible and for providing inspiring working conditions. My support from the Swedish National Board for Technical Development (STUF) has also been important.

Sabbatical or not, it was unavoidable that a lot of the writing (not to mention the thinking!) of the book had to be done on overtime. I thank my family, Ann-Kristin, Johan, and Arvid, for letting me use their time.

Lennart Ljung
Linköping, Sweden

OPERATORS AND NOTATIONAL CONVENTIONS

$\arg(z)$ = argument of the complex number z

$\arg\min f(x)$ = value of x that minimizes $f(x)$

$x_N \in AsF(n, m)$: sequence of random variables x_N converges in distribution to the F-distribution with n and m degrees of freedom

$x_N \in AsN(m, P)$: sequence of random variables x_N converges in distribution to the normal distribution with mean m and covariance matrix P; see (I.17)

$x_N \in As\chi^2(n)$: sequence of random variables x_N converges in distribution to the χ^2 distribution with n degrees of freedom

$\mathrm{Cov}(x)$ = covariance matrix of the random vector x; see (I.4)

$\det A$ = determinant of the matrix A

$\dim \theta$ = dimension (number of rows) of the column vector θ

$E\,x$ = mathematical expectation of the random vector x; see (I.3)

$\overline{E}\,x(t) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} E\,x(t)$; see (2.60)

$\mathbf{O}(x)$ = ordo x: function tending to zero at the same rate as x

$\mathbf{o}(x)$ = small ordo x: function tending to zero faster than x

$x \in N(m, P)$: random variable x is normally distributed with mean m and covariance matrix P; see (I.6)

Re z = real part of the complex number z

$\mathcal{R}(f)$ = range of the function f = the set of values that $f(x)$ may assume

$\mathbf{R}^d$ = Euclidian d-dimensional space

$x = \text{sol}\{f(x) = 0\}$: x is the solution (or set of solutions) to the equation $f(x) = 0$

$\text{tr}(A)$ = trace (the sum of the diagonal elements) of the matrix A

$\text{Var}(x)$ = variance of the random variable x

A^{-1} = inverse of the matrix A

A^T = transpose of the matrix A

A^{-T} = transpose of the inverse of the matrix A

$\bar{z}$ = complex conjugate of the complex number z

(superscript * is not used to denote transpose and complex conjugate: it is used only as a distinguishing superscript)

$y_s^t = \{y(s), y(s+1), \ldots, y(t)\}$

$y^t = \{y(1), y(2), \ldots, y(t)\}$

$U_N(\omega)$ = Fourier transform of u^N; see (2.37)

$R_v(\tau) = \bar{E}v(t)v^T(t-\tau)$; see (2.61)

$R_{sw}(\tau) = \bar{E}s(t)w^T(t-\tau)$; see (2.62)

$\Phi_v(\omega)$ = spectrum of v = Fourier transform of $R_v(\tau)$; see (2.63)

$\Phi_{sw}(\omega)$ = cross spectrum between s and w = Fourier transform of $R_{sw}(\tau)$; see (2.64)

$\hat{R}_s^N(\tau) = \frac{1}{N}\sum_{t=1}^{N} s(t)s^T(t-\tau)$; see (6.10)

$\hat{\Phi}_u^N(\omega)$ = estimate of the spectrum of u based on u^N; see (6.48)

$\hat{v}(t|t-1)$ = prediction of $v(t)$ based on v^{t-1}

$\frac{d}{d\theta}V(\theta)$ = gradient of $V(\theta)$ with respect to θ: a column vector of dimension $\dim\theta$ if V is scalar valued

$V'(\theta)$ = gradient of V with respect to its argument

$\ell'_\varepsilon(\varepsilon,\theta)$ = partial derivative of ℓ with respect to ε

δ_{ij} = Kronecker's delta: zero unless $i = j$

$\delta(k) = \delta_{k0}$

$\mathcal{B}(\theta_0,\varepsilon)$ = ε neighborhood of θ_0: $\{\theta \mid |\theta - \theta_0| < \varepsilon\}$

$\triangleq$ = the left side is defined by the right side

$|\cdot|$ = (Euclidian) norm of a vector

$\|\cdot\|$ = (Frobenius) norm of a matrix (see 2.89)

SYMBOLS USED IN TEXT

This list contains symbols that have some global use. Some of the symbols may have another local meaning.

$D_{\mathcal{M}}$ = set of values over which θ ranges in a model structure. See (4.119)

D_c = set into which the θ-estimate converges. See (8.23)

$e(t)$ = disturbance at time t; usually $\{e(t), t = 1, 2, \dots\}$ is white noise (a sequence of independent random variables with zero mean values and variance λ)

$e_0(t)$ = "true" driving disturbance acting on a given system $\mathcal{S}$; see (8.2)

$f_e(x), f_e(x,\theta)$ = probability density function of the random variable e; see (I.2) and (4.4)

$G(q)$ = transfer function from u to y; see (2.20)

$G(q,\theta)$ = transfer function in a model structure, corresponding to the parameter value θ; see (4.4)

$G_0(q)$ = "true" transfer function from u to y for a given system; see (8.7)

$\hat{G}_N(q)$ = estimate of $G(q)$ based on Z^N

$G^*(q)$ = limiting estimate of $G(q)$; see (8.68)

$\tilde{G}_N(q)$ = difference $\hat{G}_N(q) - G_0(q)$; see (8.15)

$\mathcal{G}$ = set of transfer functions obtained in a given structure; see (8.44)

$H(q), H(q,\theta), H_0(q), \hat{H}_N(q), H^*(q), \tilde{H}_N(q), \mathcal{H}$: analogous to G but for the transfer function from e to y

$L(q)$ = prefilter for the prediction errors; see (7.10)

$\ell(\varepsilon), \ell(\varepsilon,\theta), \ell(\varepsilon,t,\theta)$ = norm for the prediction errors used in the criterion; see (7.11), (7.16), (7.18)

$\mathcal{M}$ = model structure (a mapping from a parameter space to a set of models); see (4.119)

$\mathcal{M}(\theta)$ = particular model corresponding to the parameter value θ; see (4.119)

$\mathcal{M}^*$ = set of models (usually generated as the range of a model structure); see (4.115) and page 93

P_θ = asymptotic covariance matrix of θ; see (9.11)

q, q^{-1} = forward and backward shift operators; see (2.15)

$\mathcal{S}$ = "the true system"; see (8.7)

$T(q) = [G(q)\ H(q)]$; see (4.106)

$T(q,\theta), T_0(q), \hat{T}_N(q), \tilde{T}_N(q)$ = analogous to G and H

$u(t)$ = input variable at time t

$V_N(\theta, Z^N)$ = criterion function to be minimized; see (7.11)

$\bar{V}(\theta)$ = limit of criterion function; see (8.28)

$v(t)$ = disturbance variable at time t

$w(t)$ = usually a disturbance variable at time t; the precise meaning varies with the local context

$x(t)$ = state vector at time t; dimension = n

$y(t)$ = output variable at time t

$\hat{y}(t|\theta)$ = predicted output at time t using a model $\mathcal{M}(\theta)$ and based on Z^{t-1}; see (4.6)

$z(t) = [y(t)\ u(t)]^T$; see (4.110)

$Z^N = \{u(0), y(0), \ldots, u(N), y(N)\}$

$\varepsilon(t,\theta)$ = prediction error $y(t) - \hat{y}(t|\theta)$

λ = used to denote variance; also, in Chapter 11, the forgetting factor; see (11.6), (11.63)

θ = vector used to parametrize models; dimension = d; see (4.4), (4.5), (5.33)

$\hat{\theta}_N$, θ_0, θ^*, $\bar{\theta}_N$ = analogous to G

$\varphi(t)$ = regression vector at time t; see (4.11) and (5.34)

$\chi_0(t) = [u(t)\ e_0(t)]^T$; see (8.14)

$\psi(t,\theta)$ = gradient of $\hat{y}(t|\theta)$ with respect to θ; a d-dimensional column vector; see (4.118c)

$\zeta(t)$, $\zeta(t,\theta)$ = "the correlation vector" (instruments); see (7.96)

$\mathbf{T}'(q,\theta)$ = gradient of $T(q,\theta)$ with respect to θ (a $d \times 2$ matrix); see (4.122)

ABBREVIATIONS AND ACRONYMS

ARARX: See Table 4.1

ARMA: AutoRegressive Moving Average (see Table 4.1)

ARMAX: AutoRegressive Moving Average with eXternal input (see Table 4.1)

ARX: AutoRegressive with eXternal input (see Table 4.1)

BJ: Box–Jenkins model structure (see Table 4.1)

ETFE: Empirical Transfer Function Estimate; see (6.24)

FIR: Finite Impulse Response model (see Table 4.1)

IV: Instrumental variables (see Section 7.6)

LS: Least Squares (see Section 7.3)

ML: Maximum Likelihood (see Section 7.4)

MSE: Mean Square Error

OE: Output error model structure (see Table 4.1)
PDF: Probability Density Function
PEM: Prediction-Error Method (see Section 7.2)
PLR: PseudoLinear Regression (see Section 7.5)
RIV: Recursive IV (see Section 11.3)
RLS: Recursive LS (see Section 11.2)
RPEM: Recursive PEM (see Section 11.4)
RPLR: Recursive PLR (see Section 11.5)
SISO: Single Input Single Output
w.p.: with probability
w.p. 1: with probability one; see (I.15)
w.r.t.: with respect to

1

INTRODUCTION

Inferring models from observations and studying their properties is really what science is about. The models ("hypotheses," "laws of nature," "paradigms," etc.) may be of more or less formal character, but they have the basic feature that they attempt to link observations together into some pattern. System identification deals with the problem of building mathematical models of dynamical systems based on observed data from the systems. The subject is thus part of basic scientific methodology, and since dynamical systems are abundant in our environment, the techniques of system identification have a wide application area. This book aims at giving an understanding of available system identification methods, their rationale, properties, and use.

1.1 DYNAMICAL SYSTEMS

In loose terms a *system* is an object in which variables of different kinds interact and produce observable signals. The observable signals that are of interest to us are usually called *outputs*. The system is also affected by external stimuli. External signals that can be manipulated by the observer are called *inputs*. Others are called *disturbances* and can be divided into those that are directly measured and those that are only observed through their influence on the output. The distinction between inputs and measured disturbances is often less important for the modeling process. See Figure 1.1.

Clearly the notion of a system is a broad concept, and it is not surprising that it

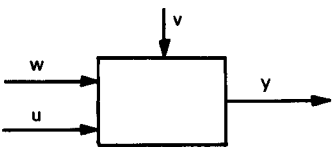

Figure 1.1 A system with output y, input u, measured disturbance w, and unmeasured disturbance v.

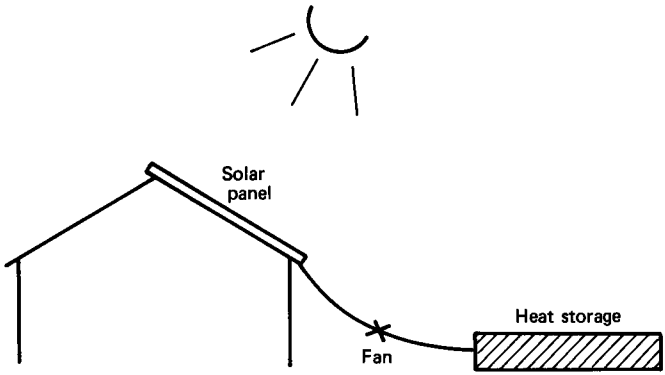

Figure 1.2 A solar-heated house.

plays an important role in modern science. Many problems in various fields are solved in a systems-oriented framework. Instead of attempting a formal definition of the system concept, we shall illustrate it by a few examples.

Example 1.1 A Solar-heated House

Consider the solar-heated house depicted in Figure 1.2. The system operates in such a way that the sun heats the air in the solar panel, which consists of transparent hoses. This air is fanned into the heat storage, which is a box filled with pebbles. The stored energy can later be transferred to the house. We are interested in how solar radiation and fan velocity affect the temperature of the heat storage. This system is symbolically depicted in Figure 1.3. Figure 1.4 shows a record of observed data over a 16-hour period. The variables were sampled every 10 minutes. ■

Example 1.2 Ship-steering Dynamics

The motion of a ship on the ocean is governed by the propeller forces, the rudder angle, and forces from wind and waves. See Figure 1.5. As a subproblem, we may consider how the ship's heading angle is affected by the rudder angle at constant propeller forces. This system is depicted in Figure 1.6. A record of data for this system is shown in Figure 1.7. It stretches over 25 minutes, and the variables were sampled every 10 seconds. ■

Example 1.3 Speech

The sound of the human voice is generated by the vibration of the vocal chords or, in the case of unvoiced sounds, the air stream from the throat, and formed by the shape of the vocal tract. See Figure 1.8. The output of this system is sound vibration (i.e., the air pressure), but

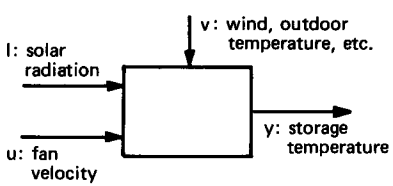

Figure 1.3 The solar-heated house system: u: input; I: measured disturbance; y: output; v: unmeasured disturbances.

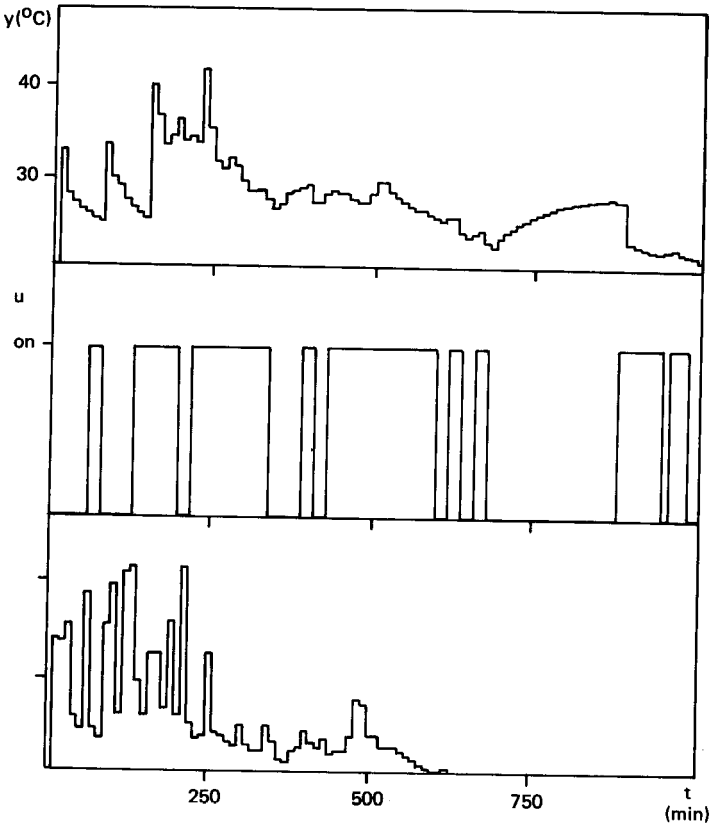

Figure 1.4 Storage temperature y, fan velocity u, and solar intensity I over a 16-hour period. Sampling interval: 10 minutes.

the external stimuli are not measurable. See Figure 1.9. Data from this system are shown in Figure 1.10. ■

The systems in these examples are all *dynamic,* which means that the current output value depends not only on the current external stimuli but also on their earlier values. Outputs of dynamical systems whose external stimuli are not observed (such as in Example 1.3) are often called *time series.* This term is especially common in economic applications. Clearly, the list of examples of dynamical systems can be very long and it stretches over many fields of science.

1.2 MODELS

Model Types and Their Use

When we interact with a system, we need some concept of how its variables relate to each other. With a broad definition, we shall call such an assumed relation-

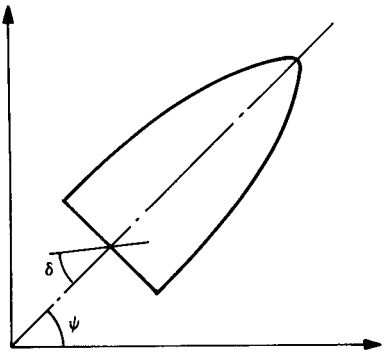

Figure 1.5 A ship's motion in the horizontal plane. δ: rudder command; ψ: heading angle.

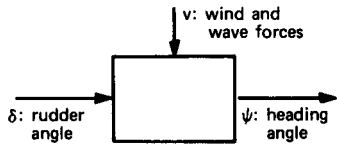

Figure 1.6 The steering dynamics system. δ: input. ψ: output. v: unmeasured disturbances.

ship among observed signals *a model* of the system. Clearly, models may come in various shapes and be phrased with varying degrees of mathematical formalism. The intended use will determine the degree of sophistication that is required to make the model purposeful.

No doubt, in daily life many systems are dealt with using *mental models*, which do not involve any mathematical formalization at all. To drive a car, for example, requires the knowledge that turning the steering wheel to the left induces a left turn, together with subtle information built up in the muscle memory. The importance and degree of sophistication of the latter should of course not be underestimated.

For certain systems it is appropriate to describe its properties using numerical tables and/or plots. We shall call such descriptions *graphical models*. Linear sys-

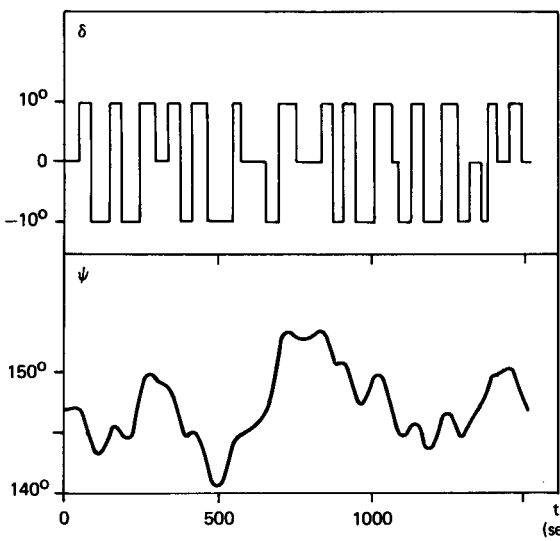

Figure 1.7 Input-output data for the ship-steering dynamics system. Sampling interval: 10 seconds.

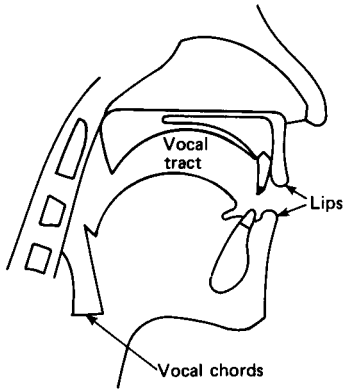

Figure 1.8 Speech generation.

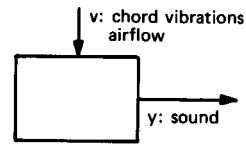

Figure 1.9 The speech system: y: output. v: unmeasured disturbance.

tems, for example, can be uniquely described by their impulse or step responses or by their frequency functions. Graphical representations of these are widely used for various design purposes. The nonlinear characteristics of, say, a valve are also well suited to be described by a graphical model.

For more advanced applications, it may be necessary to use models that describe the relationships among the system variables in terms of mathematical expressions like difference or differential equations. We shall call such models *mathematical (or analytical) models*. Mathematical models may be further characterized by a number of adjectives (time continuous or time discrete, lumped or distributed, deterministic or stochastic, linear or nonlinear, etc.) signifying the type of difference or differential equation used. The use of mathematical models is inherent in all fields of engineering and physics. In fact, a major part of the engineering field deals with how to make good designs based on mathematical models. They are also instrumental for simulation and forecasting (prediction), which is extensively used in all fields, including nontechnical areas like economy, ecology, and biology.

The model used in a computer simulation of a system is a program. For

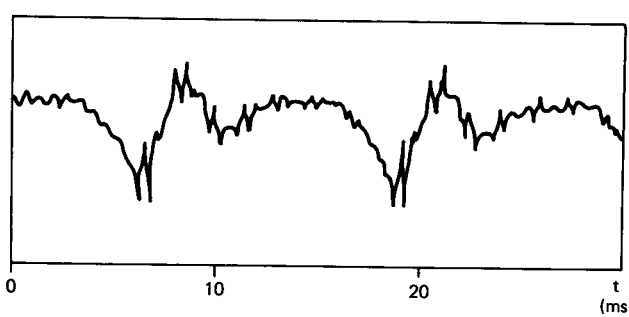

Figure 1.10 The speech signal (air pressure). Data sampled every 0.125 ms.

complex systems, this program may be built up by many interconnected subroutines and lookup tables, and it may not be feasible to summarize it analytically as a mathematical model. We use the term *software model* for such computerized descriptions. They have come to play an increasingly important role in decision making for complicated systems.

Building Models

Basically, a model has to be constructed from observed data. The mental model of car-steering dynamics, for example, is developed through driving experience. Graphical models are made up from certain measurements. Mathematical models may be developed along two routes (or a combination of them). One route is to split up the system, figuratively speaking, into subsystems, whose properties are well understood from previous experience. This basically means that we rely on "laws of nature" and other well-established relationships that have their roots in earlier empirical work. These subsystems are then joined mathematically and a model of the whole system is obtained. This route is known as *modeling* and does not necessarily involve any experimentation on the actual systems. The procedure of modeling is quite application dependent and often has its roots in tradition and specific techniques in the application area in question. Basic techniques typically involve structuring of the process into block diagrams with blocks consisting of simple elements. The reconstruction of the system from these simple blocks is now increasingly being done by computer, resulting in a software model rather than a mathematical model.

The other route to mathematical as well as graphical models is directly based on experimentation. Input and output signals from the system, such as those in Figures 1.4, 1.7, and 1.10, are recorded and subjected to data analysis in order to infer a model. This route is *system identification.*

The Fiction of a True System

The real-life actual system is an object of a different kind than our mathematical models. In a sense, there is an impenetrable but transparent screen between our world of mathematical descriptions and the real world. We can look through this window and compare certain aspects of the physical system with its mathematical description, but we can never establish any exact connection between them. The question of nature's susceptibility to mathematical description has some deep philosophical aspects, and in practical terms we have to take a more pragmatic view of models. Our acceptance of models should thus be guided by "usefulness" rather than "truth." Nevertheless, we shall occasionally use a concept of "the true system," defined in terms of a mathematical description. Such a fiction is helpful for devising identification methods and understanding their properties. In such contexts we assume that the observed data have been generated according to some well-defined mathematical rules, which of course is an idealization.

1.3 THE SYSTEM IDENTIFICATION PROCEDURE

Three Basic Entities

The construction of a model from data involves three basic entities:

1. The data
2. A set of candidate models
3. A rule by which candidate models can be assessed using the data

Let us comment on each of these:

1. *The data record.* The input–output data are sometimes recorded during a specifically designed identification experiment, where the user may determine which signals to measure and when to measure them and may also choose the input signals. The object with *experiment design* is thus to make these choices so that the data become maximally informative, subject to constraints that may be at hand. In other cases the user may not have the possibility to affect the experiment, but must use data from the normal operation of the system.
2. *The set of models.* A set of candidate models is obtained by specifying within which collection of models we are going to look for a suitable one. This is no doubt the most important and, at the same time, the most difficult choice of the system identification procedure. It is here that a priori knowledge and engineering intuition and insight have to be combined with formal properties of models. Sometimes the model set is obtained after careful *modeling*. Then a model with some unknown physical parameters is constructed from basic physical laws and other well-established relationships. In other cases standard linear models may be employed, without reference to the physical background. Such a model set, whose parameters are basically viewed as vehicles for adjusting the fit to the data and do not reflect physical considerations in the system, is called a *black box*. Model sets with adjustable parameters with physical interpretation may, accordingly, be called *gray boxes*.
3. *Determining the "best" model in the set, guided by the data.* This is the *identification method*. The assessment of model quality is typically based on how the models perform when they attempt to reproduce the measured data.

Model Validation

After having settled on the preceding three choices, we have, at least implicitly, arrived at a particular model: the one in the set that best describes the data according to the chosen criterion. It then remains to test whether this model is "good enough," that is, whether it is valid for its purpose. Such tests are known as *model validation*. They involve various procedures to assess how the model relates to observed data, to prior knowledge, and to its intended use. Deficient model

behavior in these respects make us reject the model, while good performance will develop a certain confidence in the model. A model can never be accepted as a final and true description of the system. Rather, it can at best be regarded as a good enough description of certain aspects that are of particular interest to us.

The System Identification Loop

The system identification procedure has a natural logical flow: first collect data, then choose a model set, then pick the "best" model in this set. It is quite likely, though, that the model first obtained will not pass the model validation tests. We must then go back and revise the various steps of the procedure.

The model may be deficient for a variety of reasons:

- The numerical procedure failed to find the best model according to our criterion.
- The criterion was not well chosen.
- The model set was not appropriate, in that it did not contain any "good enough" description of the system.
- The data set was not informative enough to provide guidance in selecting good models.

The major part of an identification application in fact consists of addressing these problems, in particular the third one, in an iterative manner, guided by prior information and the outcomes of previous attempts. See Figure 1.11. Interactive software obviously is an important tool for handling the iterative character of this problem.

1.4 ORGANIZATION OF THE BOOK

To master the loop of Figure 1.11, the user has to be familiar with a number of things:

1. Available techniques of identification and their rationale, as well as typical choices of model sets
2. The properties of the identified model and their dependence on the basic items: data, model set, and identification criterion
3. Numerical schemes for computing the estimate
4. How to make intelligent choices of experiment design, model set, and identification criterion, guided by prior information as well as by observed data

In fact, a user of system identification may find that he or she is primarily a user of an interactive identification software package. Items 1 and 3 are then part of

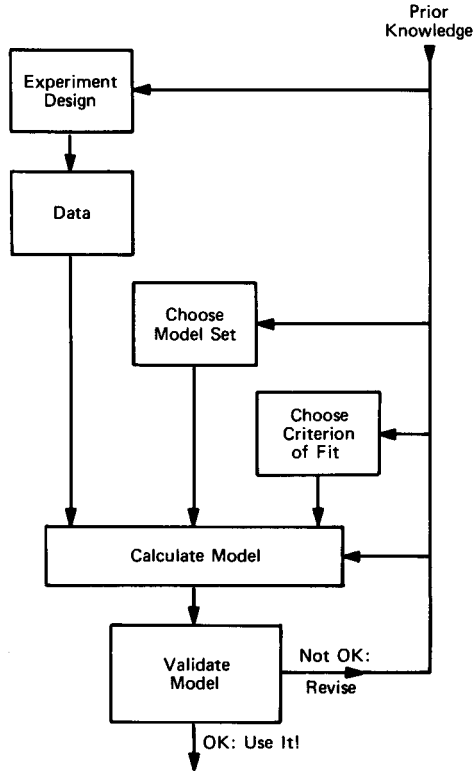

Figure 1.11 The system identification loop.

the package and the important thing is to have a good understanding of item 2 so that task 4 can be successfully completed. This is what we mean by "Theory for the User," and this is where the present book has its focus.

The idea behind the book's organization is to present the list of common and useful model sets in Chapters 4 and 5. Available techniques are presented in Chapters 6 and 7, and the analysis follows in Chapters 8 and 9. Numerical techniques for off-line and on-line applications are described in Chapters 10 and 11. Task 4, the user's choices, is discussed primarily in Chapters 14 through 16, after some preliminaries in Chapter 12 and 13. In addition, Chapters 2 and 3 give the formal setup of the book, and Chapter 17 describes and assesses system identification as a tool for practical problems.

Figure 1.12 illustrates the book's structure in relation to the loop of system identification. A flow chart of the chapters of the book is given in Figure 1.13. It shows, among other things, which chapters can be skipped without serious loss of continuity.

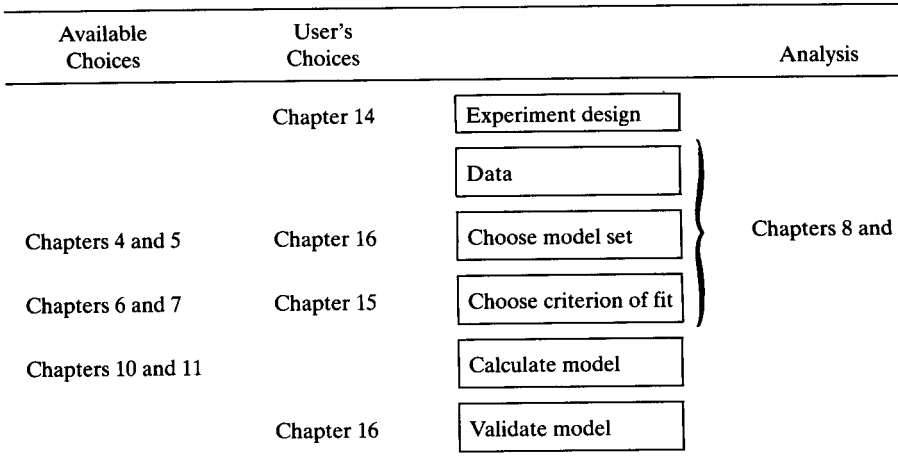

Figure 1.12 Organization of the book.

About the Framework

The system identification framework we set up here is fairly general. It does not confine us to linear models or quadratic criteria or to assuming that the system itself can be described within the model set. Indeed, this is one of the points that should be stressed about our framework. Nevertheless, we often give proofs and explicit expressions only for certain special cases, like single-input, single-output systems and quadratic criteria. The purpose is of course to enhance the underlying basic ideas and not conceal them behind technical details. References are usually provided for more general treatments.

Parameter estimation and identification are usually described within a probabilistic framework. Here we basically employ such a framework. However, we also try to retain a pragmatic viewpoint that is independent of probabilistic interpretations. That is, the methods we describe and the recommendations we put forward should make sense even without the probabilistic framework that may motivate them as "optimal solutions." The probabilistic and statistical environments of the book are described in Appendixes I and II, respectively. These appendixes may be read prior to the other chapters or consulted occasionally when required. In any case, the book does not lean heavily on the background provided there.

1.5 BIBLIOGRAPHY

The literature on the system identification problem and its ramifications is extensive. Among general textbooks on the subject we may mention Box and Jenkins (1970), Eykhoff (1974), and Spriet and Vansteenkiste (1982) for thorough treatments covering several practical issues, while Goodwin and Payne (1977), Davis and Vinter (1985), and Söderström and Stoica (1987) give more theoretically oriented

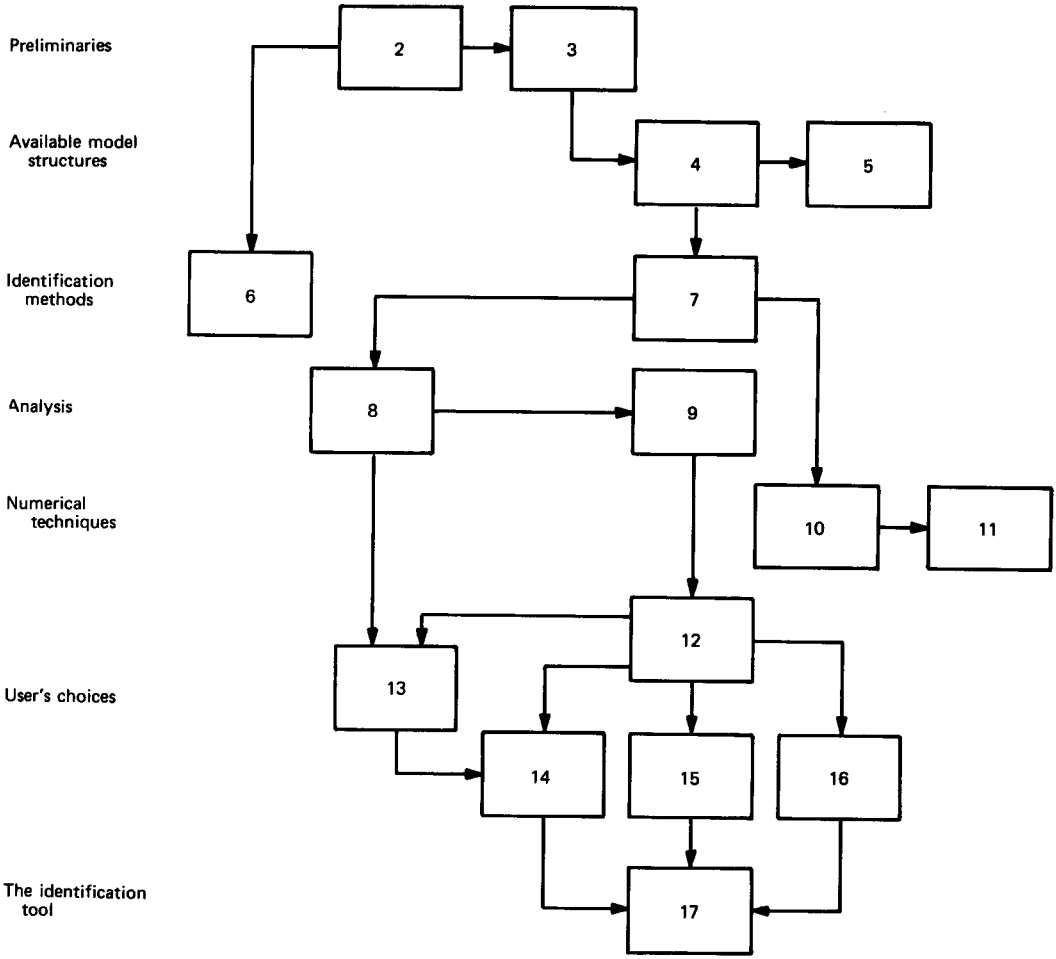

Figure 1.13 Flowchart of the book's chapters.

presentations. Kashyap and Rao (1976) emphasize the role of model validation and model selection in their treatment of system identification, while Söderström and Stoica (1983) focus on instrumental-variable methods. Texts that concentrate on recursive identification techniques include Ljung and Söderström (1983) and Young (1984). Statistical treatments of time-series modeling, such as T. Anderson (1971), Hannan (1970), and Brillinger (1981), are most relevant also for the system identification problem.

Among edited collections of articles, we may refer to Mehra and Lainiotis (1976), Eykhoff (1981), Hannan, Krishnaiah, and Rao (1985), and Leondes (1987), as well as to the special journal issues Kailath, Mehra, and Mayne (1974) and Isermann (1981). The proceedings from the International Federation of Automatic Control (IFAC) series, Symposium on Identification and System Parameter

Estimation (Prague, 1967, 1970; the Hague, 1973; Tbilisi, 1976; Darmstadt, 1979; Washington, D.C., 1982; York, 1985; Beijing, 1988), contain many articles on all aspects of the system identification problem.†

Philosophical aspects on mathematical models of real-life objects are discussed, for example, in Popper (1934). Modeling from basic physical laws, rather than from data, is discussed in many books; see, for example, Wellstead (1979), Nicholson (1981), and Frederick and Close (1978) for engineering applications. Such treatments are important complements to the model set selection (see Section 1.3 and Chapter 16).

Many books discuss modeling and identification in various application areas. See for example, Granger and Newbold (1977) or Malinvaud (1980) (econometrics), Godfrey (1983) (biology), Robinson and Treitel (1980) or Mendel (1983) (geoscience), Dudley (1983) (electromagnetic wave theory), Markel and Gray (1976) (speech signals), and Beck and van Straten (1983) (environmental systems). Rajbman (1975) has surveyed the Soviet literature.

† The proceedings of the Hague and Tbilisi Symposia have been published by North-Holland, Amsterdam, while the proceedings of the subsequent symposia have been published by Pergamon Press, New York.

I systems and models

2

TIME-INVARIANT LINEAR SYSTEMS

Time-invariant linear systems no doubt form the most important class of dynamical systems considered in practice and in the literature. It is true that they represent idealizations of the processes encountered in real life. But, even so, the approximations involved are often justified, and design considerations based on linear theory lead to good results in many cases.

A treatise of linear systems theory is a standard ingredient in basic engineering education, and the reader has no doubt some knowledge of this topic. Anyway, in this chapter we shall provide a refresher on some basic concepts that will be instrumental for the further development in this book. In Section 2.1 we shall discuss the impulse response and various ways of describing and understanding disturbances, as well as introduce the transfer-function concept. In Section 2.2 we study frequency-domain interpretations and also introduce the periodogram. Section 2.3 gives a unified setup of spectra of deterministic and stochastic signals that will be used in the remainder of this book. In Section 2.4 a basic ergodicity result is proved. The development in these sections is for systems with a scalar input and a scalar output. Section 2.5 contains the corresponding expressions for multivariable systems.

2.1 IMPULSE RESPONSES, DISTURBANCES, AND TRANSFER FUNCTIONS

Impulse Response

Consider a system with a scalar input signal $u(t)$ and a scalar output signal $y(t)$ (Figure 2.1). The system is said to be *time invariant* if its response to a certain input signal does not depend on absolute time. It is said to be *linear* if its output response

Figure 2.1 The system

to a linear combination of inputs is the same linear combination of the output responses of the individual inputs. Furthermore, it is said to be *causal* if the output at a certain time depends on the input up to that time only.

It is well known that a linear, time-invariant, causal system can be described by its *impulse response* (or *weighting function*) $g(\tau)$ as follows:

$$y(t) = \int_{\tau=0}^{\infty} g(\tau)u(t-\tau)d\tau \tag{2.1}$$

Knowing $\{g(\tau)\}_{\tau=0}^{\infty}$ and knowing $u(s)$ for $s \leq t$, we can consequently compute the corresponding output $y(s)$, $s \leq t$ for any input. *The impulse response is thus a complete characterization of the system.*

Sampling

In this book we shall almost exclusively deal with observations of inputs and outputs in discrete time, since this is the typical data-acquisition mode. We thus assume $y(t)$ to be observed at the *sampling instants* $t_k = kT$, $k = 1,2,\ldots$:

$$y(kT) = \int_{\tau=0}^{\infty} g(\tau)\, u(kT-\tau)\, d\tau \tag{2.2}$$

The interval T will be called the *sampling interval.* It is, of course, also possible to consider the situation where the sampling instants are not equally spread.

Most often, in computer control applications, the input signal $u(t)$ is kept constant between the sampling instants:

$$u(t) = u_k, \qquad kT \leq t < (k+1)T \tag{2.3}$$

This is mostly done for practical implementation reasons, but it will also greatly simplify the analysis of the system. Inserting (2.3) into (2.2) gives

$$\begin{aligned} y(kT) &= \int_{\tau=0}^{\infty} g(\tau)\, u(kT-\tau)\, d\tau = \sum_{\ell=1}^{\infty} \int_{\tau=(\ell-1)T}^{\ell T} g(\tau)\, u(kT-\tau)\, d\tau \\ &= \sum_{\ell=1}^{\infty} \left[\int_{\tau=(\ell-1)T}^{\ell T} g(\tau)\, d\tau \right] u_{k-\ell} = \sum_{\ell=1}^{\infty} g_T(\ell)\, u_{k-\ell} \end{aligned} \tag{2.4}$$

where we defined

$$g_T(\ell) = \int_{\tau=(\ell-1)T}^{\ell T} g(\tau)d\tau \tag{2.5}$$

The expression (2.4) tells us what the output will be at the sampling instants. Note that no approximation is involved if the input is subject to (2.3) and that it is

sufficient to know the sequence $\{g_T(\ell)\}_{\ell=1}^{\infty}$ in order to compute the response to the input. The relationship (2.4) describes a *sampled-data system,* and we shall call the sequence $\{g_T(\ell)\}_{\ell=1}^{\infty}$ the impulse response of that system.

Even if the input is not piecewise constant and subject to (2.3), the representation (2.4) might still be a reasonable approximation, provided $u(t)$ does not change too much during a sampling interval. See also the following expressions (2.21) to (2.26).

We shall stick to the notation (2.3) to (2.5) when the choice and size of T are essential to the discussion. For most of the time, however, we shall for ease of notation assume that T is one time unit and use t to enumerate the sampling instants. We thus write for (2.4).

$$y(t) = \sum_{k=1}^{\infty} g(k)u(t-k), \qquad t = 0,1,2,\ldots \tag{2.6}$$

For sequences, we shall also use the notation

$$y_s^t = (y(s), y(s+1), \ldots, y(t)) \tag{2.7}$$

and for simplicity

$$y_1^t = y^t$$

Disturbances

According to the relationship (2.6), the output can be exactly calculated once the input is known. In most cases this is unrealistic. There are always signals beyond our control that also affect the system. Within our linear framework we assume that such effects can be lumped into an additive term $v(t)$ at the output (see Figure 2.2):

$$y(t) = \sum_{k=1}^{\infty} g(k)u(t-k) + v(t) \tag{2.8}$$

There are many sources and causes for such a disturbance term. We could list

- Measurement noise: The sensors that measure the signals are subject to noise and drift.
- Uncontrollable inputs: The system is subject to signals that have the character of inputs, but are not controllable by the user. Think of an airplane, whose movements are affected by the inputs of rudder and aileron deflections, but also by wind gusts and turbulence. Another example could be a room, where the temperature is determined by radiators, whose effect we control, but also by people (≈100 W per person) who may move in and out in an unpredictable manner.

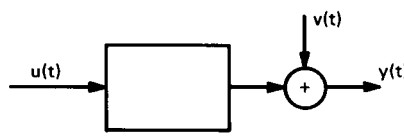

Figure 2.2 System with disturbance.

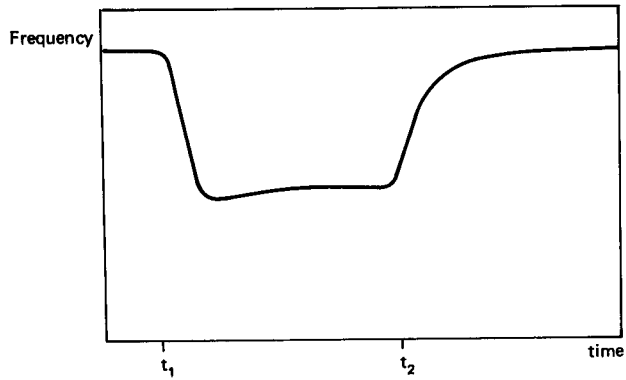

Figure 2.3 Frequency of generated ac voltage, subject to a load change between times t_1 and t_2.

The character of the disturbances could also vary within wide ranges. Classical ways of describing disturbances in control have been to study steps and pulses, while in stochastic control the disturbances are modeled as realizations of stochastic processes. See Figures 2.3 and 2.4 for some typical, but mutually quite different, disturbance characteristics. The disturbances may in some cases be separately measurable, but in the typical situation they are noticeable only via their effect on the output. If the impulse response of the system is known, then of course the actual value of the disturbance $v(t)$ can be calculated from (2.8) at time t.

The assumption of Figure 2.2 that the noise enters additively to the output implies some restrictions. Sometimes the measurements of the inputs to the system may also be noise corrupted ("error-in-variable" descriptions). In such cases we take a pragmatic approach and regard the measured input values as the actual inputs $u(t)$ to the process, and their deviations from the true stimuli will be propagated through the system and lumped into the disturbance $v(t)$ of Figure 2.2.

Characterization of Disturbances

The most characteristic feature of a disturbance is that *its value is not known beforehand.* Information about past disturbances could, however, be important for

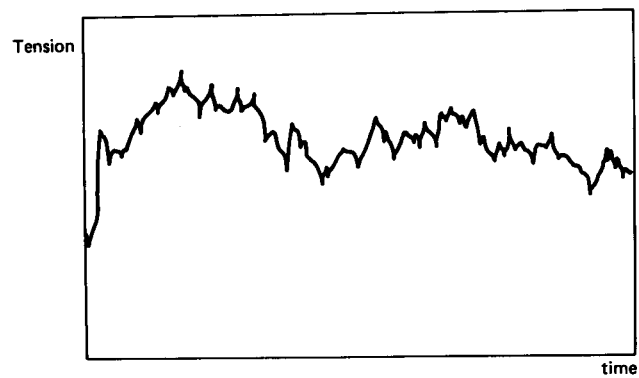

Figure 2.4 Tension of paper in the dryer part of a paper machine.

making qualified guesses about coming values. It is thus natural to employ a probabilistic framework to describe future disturbances. We then put ourselves at time t and would like to make a statement about disturbances at times $t + k$, $k \geq 1$. A complete characterization would be to describe the conditional joint probability density function for $\{v(t + k),\ k \geq 1\}$, given $\{v(s),\ s \leq t\}$. This would, however, in most cases be too laborious, and we shall instead use a simpler approach.

Let $v(t)$ be given as

$$v(t) = \sum_{k=0}^{\infty} h(k)e(t - k) \tag{2.9}$$

where $\{e(t)\}$ is a sequence of independent (identically distributed) random variables with a certain probability density function. Although this description does not allow completely general characterizations of all possible probabilistic disturbances, it is versatile enough for most practical purposes. In Section 3.1 we shall show how the description (2.9) allows predictions and probabilistic statements about future disturbances. For normalization reasons, we shall usually assume that $h(0) = 1$, which is no loss of generality since the variance of e can be adjusted.

It should be made clear that the specification of different probability density functions (PDF) for $\{e(t)\}$ may result in very different characteristic features of the disturbance. For example, the PDF

$$\begin{aligned} e(t) &= 0, \qquad \text{with probability } 1 - \mu \\ e(t) &= r, \qquad \text{with probability } \mu \end{aligned} \tag{2.10}$$

where r is a normally distributed random variable: $r \in N(0,\gamma)$ leads to, if μ is a small number, disturbance sequences with characteristic and "deterministic" profiles occurring at random instants. See Figure 2.5. This could be suitable to describe "classical" disturbance patterns, steps, pulses, sinusoids, and ramps (cf. Figure 2.3!). On the other hand, the PDF

$$e(t) \in N(0,\lambda) \tag{2.11}$$

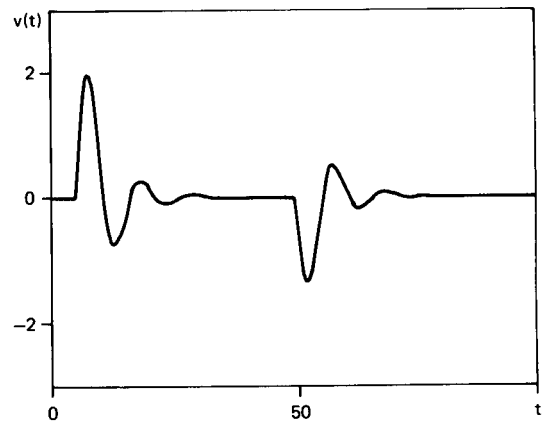

Figure 2.5 A realization of the process (2.9) with e subject to (2.10).

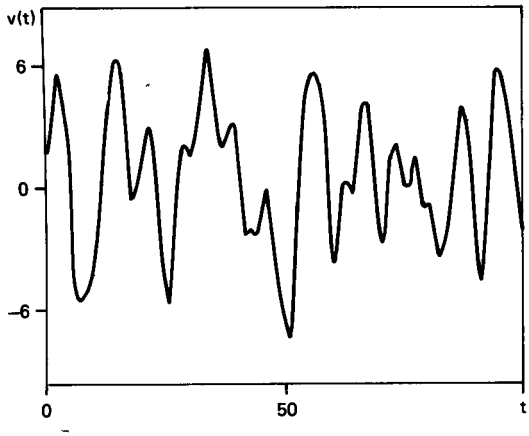

Figure 2.6 A realization of the same process (2.9) as in Figure 2.5, but with e subject to (2.11).

gives a totally different picture. See Figure 2.6. Such a pattern is more suited to describe measurements noises and irregular and frequent disturbance actions of the "input type."

Often we only specify the *second-order properties* of the sequence $\{e(t)\}$, that is, the mean and the variances. Note that (2.10) and (2.11) can both be described as "a sequence of independent random variables with zero mean values and variances λ" [$\lambda = \mu\gamma$ for (2.10)], despite the difference in appearance.

Remark. Notice that $\{e(t)\}$ and $\{v(t)\}$ as defined previously are *stochastic processes* (i.e., sequences of random variables). The disturbances that we observe and that are added to the system output as in Figure 2.2 are thus *realizations* of the *stochastic process* $\{v(t)\}$. Strictly speaking, one should distinguish in notation between the process and its realization, but the meaning is usually clear from the context, and we do not here adopt this extra notational burden. Often one has occasion to study signals that are mixtures of deterministic and stochastic components. A framework for this will be discussed in Section 2.3.

Covariance Function

With the description (2.9) of $v(t)$, we can compute the mean as

$$Ev(t) = \sum_{k=0}^{\infty} h(k)Ee(t-k) = 0 \tag{2.12}$$

and the covariance as

$$\begin{aligned} Ev(t)v(t-\tau) &= \sum_{k=0}^{\infty}\sum_{s=0}^{\infty} h(k)h(s)Ee(t-k)e(t-\tau-s) \\ &= \sum_{k=0}^{\infty}\sum_{s=0}^{\infty} h(k)h(s)\delta(k-\tau-s)\lambda \\ &= \lambda\sum_{k=0}^{\infty} h(k)h(k-\tau) \end{aligned} \tag{2.13}$$

Here $h(r) = 0$ if $r < 0$. We note that this covariance is independent of t and call

$$R_v(\tau) = Ev(t)v(t - \tau) \tag{2.14}$$

the *covariance function of the process v*. This function, together with the mean, specifies the *second-order properties of v*. These are consequently uniquely defined by the sequence $\{h(k)\}$ and the variance λ of e. Since (2.14) and $Ev(t)$ do not depend on t, the process is said to be *stationary*.

Transfer Functions

It will be convenient to introduce a shorthand notation for sums like (2.8) and (2.9), which will occur frequently in this book. We introduce the *forward shift operator q* by

$$qu(t) = u(t + 1)$$

and the *backward shift operator* q^{-1}:

$$q^{-1}u(t) = u(t - 1)$$

We can then write for (2.6)

$$\begin{aligned} y(t) &= \sum_{k=1}^{\infty} g(k)u(t - k) = \sum_{k=1}^{\infty} g(k)(q^{-k}u(t)) \\ &= \left[\sum_{k=1}^{\infty} g(k)q^{-k}\right]u(t) = G(q)u(t) \end{aligned} \tag{2.15}$$

where we introduced the notation

$$G(q) = \sum_{k=1}^{\infty} g(k)q^{-k} \tag{2.16}$$

We shall call $G(q)$ the *transfer operator* or the *transfer function* of the linear system (2.6). Notice that (2.15) thus describes a relation between the *sequences* u^t and y^t.

Remark. We choose q as argument of G rather than q^{-1} (which perhaps would be more natural in view of the right side) in order to be in formal agreement with z-transform and Fourier-transform expression. Strictly speaking, the term transfer *function* should be reserved for the z-transform of $\{g(k)\}_1^\infty$, that is,

$$G(z) = \sum_{k=1}^{\infty} g(k)z^{-k} \tag{2.17}$$

but we shall sometimes not observe that point. ■

Similarly with

$$H(q) = \sum_{k=0}^{\infty} h(k)q^{-k} \tag{2.18}$$

we can write

$$v(t) = H(q)e(t) \tag{2.19}$$

for (2.9). Our basic description for a linear system with additive disturbance will thus be

$$y(t) = G(q)u(t) + H(q)e(t) \tag{2.20}$$

with $\{e(t)\}$ as a sequence of independent random variables with zero mean values and variances λ.

Continuous-time Representation and Sampling Transfer Functions (*)

For many physical systems it is natural to work with a continuous-time representation (2.1), since most basic relationships are expressed in terms of differential equations. With $G_c(s)$ denoting the Laplace transform of the impulse response function $\{g(\tau)\}$ in (2.1), we then have the relationship

$$Y(s) = G_c(s)U(s) \tag{2.21}$$

between $Y(s)$ and $U(s)$, the Laplace transforms of the output and input, respectively. Introducing p as the differentiation operator, we could then write

$$y(t) = G_c(p)u(t) \tag{2.22}$$

as a shorthand operator form of (2.1) or its underlying differential equation. Now, (2.1) or (2.22) describes the output at all values of the continuous time variable t. If $\{u(t)\}$ is a known function (piecewise constant or not), then (2.22) will of course also serve as a description of the output of the sampling instants. We shall therefore occasionally use (2.22) also as a system description for the sampled output values, keeping in mind that the computation of these values will involve numerical solution of a differential equation. In fact, we could still use a discrete-time model (2.9) for the disturbances that influence our discrete-time measurements, writing this as

$$y(t) = G_c(p)u(t) + H(q)e(t), \qquad t = 1,2,\ldots \tag{2.23}$$

Often, however, we shall go from the continuous-time representation (2.22) to the standard discrete-time one (2.15) by transforming the transfer function

$$G_c(p) \rightarrow G_T(q) \tag{2.24}$$

T here denotes the sampling interval. When the input is piecewise constant over the sampling interval, this can be done without approximation, in view of (2.4). See Problem 2G.4 for a direct transfer-function expression, and equations (4.65) to (4.68), for numerically more favorable expressions. One can also apply approximate formulas that correspond to replacing the differentiation operator p by a difference approximation. We thus have the Euler approximation

$$G_T(q) \approx G_c\left(\frac{q-1}{T}\right) \tag{2.25}$$

(*)Denotes sections and subsections that are optional reading; they can be omitted without serious loss of continuity. See Preface.

and Tustin's formula

$$G_T(q) \approx G_c\left(\frac{2}{T}\frac{q-1}{1+q}\right) \tag{2.26}$$

See Åström and Wittenmark (1984) for a further discussion.

Some Terminology

The function $G(z)$ in (2.17) is a complex-valued function of the complex variable z. Values β_i, such that $G(\beta_i) = 0$, are called *zeros* of the transfer function (or of the system), while values α_i for which $G(z)$ tends to infinity are called *poles*. This coincides with the terminology for analytic functions (see, e.g., Ahlfors, 1979). If $G(z)$ is a rational function of z, the poles will be the zeros of the denominator polynomial.

We shall say that the transfer function $G(q)$ (or "the system G" or "the filter G") is *stable* if

$$G(q) = \sum_{k=1}^{\infty} g(k)q^{-k}, \qquad \sum_{k=1}^{\infty} |g(k)| < \infty \tag{2.27}$$

The definition (2.27) coincides with the system theoretic definition of bounded-input, bounded-output (BIBO) stability (e.g. Brockett, 1970): If an input $\{u(t)\}$ to $G(q)$ is subject to $|u(t)| \le C$, then the corresponding output $z(t) = G(q)u(t)$ will also be bounded, $|z(t)| \le C'$, provided (2.27) holds. Notice also that (2.27) assures that the (Laurent) expansion

$$G(z) = \sum_{k=1}^{\infty} g(k)z^{-k}$$

is convergent for all $|z| \ge 1$. This means that the function $G(z)$ *is analytic on and outside the unit circle.* In particular, it then has no poles in that area.

We shall often have occasion to consider families of filters $G_\alpha(q)$, $\alpha \in \mathcal{A}$:

$$G_\alpha(q) = \sum_{k=1}^{\infty} g_\alpha(k)q^{-k}, \qquad \alpha \in \mathcal{A} \tag{2.28}$$

We shall then say that such *a family is uniformly stable* if

$$|g_\alpha(k)| \le g(k), \forall\, \alpha \in \mathcal{A}, \sum_{k=1}^{\infty} g(k) < \infty \tag{2.29}$$

Sometimes a slightly stronger condition than (2.27) will be required. We shall say that $G(q)$ is *strictly stable* if

$$\sum_{k=1}^{\infty} k|g(k)| < \infty \tag{2.30}$$

Notice that, for a transfer function that is rational in q, stability implies strict stability (and, of course, vice versa). See Problem 2T.3.

Finally, we shall say that a filter $H(q)$ is *monic* if its zeroth coefficient is 1 (or the unit matrix):

$$H(q) = \sum_{k=0}^{\infty} h(k)q^{-k}, \qquad h_0 = 1 \tag{2.31}$$

2.2 FREQUENCY-DOMAIN EXPRESSIONS

Sinusoid Response and the Frequency Function

Suppose that the input to the system (2.6) is a sinusoid:

$$u(t) = \cos \omega t \tag{2.32}$$

It will be convenient to rewrite this as

$$u(t) = \mathrm{Re}\, e^{i\omega t}$$

with Re denoting "real part." According to (2.6), the corresponding output will be

$$\begin{aligned} y(t) &= \sum_{k=1}^{\infty} g(k)\,\mathrm{Re}\; e^{i\omega(t-k)} = \mathrm{Re}\sum_{k=1}^{\infty} g(k)e^{i\omega(t-k)} \\ &= \mathrm{Re}\left\{e^{i\omega t}\cdot \sum_{k=1}^{\infty} g(k)\, e^{-i\omega k}\right\} = \mathrm{Re}\{e^{i\omega t}\cdot G(e^{i\omega})\} \\ &= |G(e^{i\omega})|\cos(\omega t + \varphi) \end{aligned} \tag{2.33}$$

where

$$\varphi = \arg G(e^{i\omega}) \tag{2.34}$$

Here, the second equality follows since the $g(k)$ are real and the fourth equality from the definition (2.16) or (2.17). The fifth equality follows straightforward rules for complex numbers.

In (2.33) we assumed that the input was a cosine since time minus infinity. If $u(t) = 0, t < 0$, we obtain an additional term

$$-\mathrm{Re}\{e^{i\omega t}\sum_{k=t}^{\infty} g(k)e^{-i\omega k}\}$$

in (2.33). This term is dominated by

$$\sum_{k=t}^{\infty} |g(k)|$$

and therefore is of transient nature (tends to zero as t tends to infinity), provided that $G(q)$ is stable.

In any case, (2.33) tells us that the output to (2.32) will also be a cosine of the

same frequency, but with an amplitude magnified by $|G(e^{i\omega})|$ and a phase shift of $\arg G(e^{i\omega})$ radians. The complex number

$$G(e^{i\omega}) \tag{2.35}$$

which is the transfer function evaluated at the point $z = e^{i\omega}$, therefore gives full information as to what will happen in stationarity, when the input is a sinusoid of frequency ω. For that reason, the complex-valued function

$$G(e^{i\omega}), \qquad -\pi \le \omega \le \pi \tag{2.36}$$

is called the *frequency function* of the system (2.6). It is customary to graphically display this function as $\log|G(e^{i\omega})|$ and $\arg G(e^{i\omega})$ plotted against $\log \omega$ in a *Bode plot.* The plot of (2.36) in the complex plane is called the *Nyquist plot.* These concepts are probably better known in the continuous-time case, but all their basic properties carry over to the sampled-data case.

Periodograms of Signals over Finite Intervals

Consider the finite sequence of inputs $u(t)$, $t = 1,2,\ldots,N$. Let us define the function $U_N(\omega)$ by

$$U_N(\omega) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} u(t)\, e^{-i\omega t} \tag{2.37}$$

The values obtained for $\omega = 2\pi k/N$, $k = 1,\ldots,N$, form the familiar discrete Fourier transform (DFT) of the sequence u_1^N. We can then represent $u(t)$ by the inverse DFT as

$$u(t) = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} U_N\left(\frac{2\pi k}{N}\right) e^{i2\pi kt/N} \tag{2.38}$$

To prove this, we insert (2.37) into the right side of (2.38), giving

$$\frac{1}{N} \sum_{k=1}^{N} \sum_{s=1}^{N} u(s) \exp\left(-\frac{i2\pi ks}{N}\right) \cdot \exp\left(\frac{i2\pi kt}{N}\right)$$

$$= \frac{1}{N} \sum_{s=1}^{N} u(s) \sum_{k=1}^{N} \exp\left(\frac{2\pi ik(t-s)}{N}\right) = \frac{1}{N} \sum_{s=1}^{N} u(s) N \cdot \delta(t-s) = u(t)$$

Here we used the relationship

$$\frac{1}{N} \sum_{k=1}^{N} e^{2\pi irk/N} = \begin{cases} 1, & r = 0 \\ 0, & 1 \le r < N \end{cases} \tag{2.39}$$

From (2.37) we note that $U_N(\omega)$ is periodic with period 2π:

$$U_N(\omega + 2\pi) = U_N(\omega) \tag{2.40}$$

Also, since $u(t)$ is real,

$$U_N(-\omega) = \overline{U_N(\omega)} \tag{2.41}$$

where the overbar denotes the complex conjugate. The function $U_N(\omega)$ is therefore uniquely defined by its values over the interval $[0,\pi]$. It is, however, customary to consider $U_N(\omega)$ for $-\pi \le \omega \le \pi$, and in accordance with this (2.38) is usually written

$$u(t) = \frac{1}{\sqrt{N}} \sum_{k=-N/2+1}^{N/2} U_N\left(\frac{2\pi k}{N}\right) e^{i2\pi kt/N} \tag{2.42}$$

making use of (2.40) and the periodicity of $e^{i\omega}$. In (2.42) and elsewhere we assume N to be even; for odd N analogous summation boundaries apply.

In (2.42) we represent the signal $u(t)$ as a linear combination of $e^{i\omega t}$ for N different frequencies ω. As is further elaborated in Problem 2D.1, this can also be rewritten as sums of $\cos \omega t$ and $\sin \omega t$ for the same frequencies, thus avoiding complex numbers.

The number $U_N(2\pi k/N)$ tells us the "weight" that the frequency $\omega = 2\pi k/N$ carries in the decomposition of $\{u(t)\}_{t=1}^{N}$. Its absolute square value $|U_N(2\pi k/N)|^2$ is therefore a measure of the energy contribution of this frequency to the "signal effect." This value

$$|U_N(\omega)|^2 \tag{2.43}$$

is known as the *periodogram* of the signal $u(t)$, $t = 1,2,\ldots,N$.

Parseval's relationship,

$$\sum_{k=1}^{N} \left| U_N\left(\frac{2\pi k}{N}\right) \right|^2 = \sum_{t=1}^{N} u^2(t) \tag{2.44}$$

reinforces the interpretation that the energy of the signal can be decomposed into energy contributions from different frequencies. Think of the analog decomposition of light into its spectral components!

Example 2.1 Periodogram of a Sinusoid

Suppose that

$$u(t) = A \cos \omega_0 t \tag{2.45}$$

where $\omega_0 = 2\pi/N_0$ for some integer $N_0 > 1$. Consider the interval $t = 1,2,\ldots,N$, where N is a multiple of N_0: $N = s \cdot N_0$. Writing

$$\cos \omega_0 t = \tfrac{1}{2}[e^{i\omega_0 t} + e^{-i\omega_0 t}]$$

gives

$$U_N(\omega) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} \frac{A}{2}[e^{i(\omega_0 - \omega)t} + e^{-i(\omega_0 + \omega)t}]$$

Using (2.39), we find that

$$|U_N(\omega)|^2 = \begin{cases} N \cdot \dfrac{A^2}{4}, & \text{if } \omega = \pm\omega_0 = \dfrac{2\pi}{N_0} = \dfrac{2\pi s}{N} \\ 0, & \text{if } \omega = \dfrac{2\pi k}{N}, \quad k \neq s \end{cases} \tag{2.46}$$

The periodogram thus has two spikes in the interval $[-\pi, \pi]$. ■

Example 2.2 Periodogram of a Periodic Signal

Suppose $u(t) = u(t + N_0)$ and we consider the signal over the interval $[1, N]$, $N = s \cdot N_0$. According to (2.42), the signal over the interval $[1, N_0]$ can be written

$$u(t) = \frac{1}{\sqrt{N_0}} \sum_{r=-N_0/2+1}^{N_0/2} A_r e^{2\pi i t r/N_0} \tag{2.47}$$

with

$$A_r = \frac{1}{\sqrt{N_0}} \sum_{t=1}^{N_0} u(t) e^{-2\pi i t r/N_0} \tag{2.48}$$

Since u is periodic, (2.47) applies over the whole interval $[1,N]$. It is thus a sum of N_0 sinusoids, and the results of the previous example (or straightforward calculations) show that

$$|U_N(\omega)|^2 = \begin{cases} s \cdot |A_r|^2, & \text{if } \omega = \dfrac{2\pi r}{N_0}, \quad r = 0, \pm 1 \pm \cdots \pm \dfrac{N_0}{2} \\ 0, & \text{if } \omega = \dfrac{2\pi k}{N}, \quad k \neq r \cdot s \end{cases} \tag{2.49}$$

■

The periodograms of Examples 2.1 and 2.2 turned out to be well behaved. For signals that are realizations of stochastic processes, the periodogram is typically a very erratic function of frequency. See Figure 2.8 and Lemma 6.2.

Transformation of Periodograms(*)

As a signal is filtered through a linear system, its periodogram changes. We show next how a signal's Fourier transform is affected by linear filtering. Results for the transformation of periodograms are then immediate.

Theorem 2.1. Let $\{s(t)\}$ and $\{w(t)\}$ be related by the strictly stable system $G(q)$:

$$s(t) = G(q)w(t) \tag{2.50}$$

The input $w(t)$ for $t \leq 0$ is unknown, but obeys $|w(t)| \leq C_W$ for all t. Let

$$S_N(\omega) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} s(t) e^{-i\omega t} \tag{2.51}$$

$$W_N(\omega) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} w(t) e^{-i\omega t} \tag{2.52}$$

Then

$$S_N(\omega) = G(e^{i\omega}) W_N(\omega) + R_N(\omega) \tag{2.53}$$

where

$$|R_N(\omega)| \le 2C_w \cdot \frac{C_G}{\sqrt{N}} \tag{2.54}$$

with

$$C_G = \sum_{k=1}^{\infty} k|g(k)| \tag{2.55}$$

Proof. We have by definition

$$S_N(\omega) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} s(t)e^{-it\omega} = \frac{1}{\sqrt{N}} \sum_{k=1}^{\infty} \sum_{t=1}^{N} g(k)w(t-k)e^{-it\omega}$$

$$= [\text{change variables: } t-k=\tau] = \frac{1}{\sqrt{N}} \sum_{k=1}^{\infty} g(k)e^{-ik\omega} \cdot \sum_{\tau=1-k}^{N-k} w(\tau)e^{-i\tau\omega}$$

Now

$$\left|W_N(\omega) - \frac{1}{\sqrt{N}} \sum_{\tau=1-k}^{N-k} w(\tau)e^{-i\tau\omega}\right|$$

$$\le \left|\frac{1}{\sqrt{N}} \sum_{\tau=1-k}^{0} w(\tau)e^{-i\tau\omega}\right| + \left|\frac{1}{\sqrt{N}} \sum_{N-k+1}^{N} w(\tau)e^{-i\tau\omega}\right| \le \frac{2}{\sqrt{N}} \cdot k \cdot C_w \tag{2.56}$$

Hence

$$|S_N(\omega) - G(e^{i\omega})W_N(\omega)| = \left|\sum_{k=1}^{\infty} g(k)e^{-ik\omega}\left[\frac{1}{\sqrt{N}} \sum_{\tau=1-k}^{N-k} w(\tau)e^{-i\tau\omega} - W_N(\omega)\right]\right|$$

$$\le \frac{2}{\sqrt{N}} \sum_{k=1}^{\infty} |k \cdot g(k)C_w e^{-ik\omega}| \le \frac{2C_w \cdot C_G}{\sqrt{N}}$$

and (2.53) to (2.55) follow. ■

Corollary. Suppose $\{w(t)\}$ is periodic with period N. Then $R_N(\omega)$ in (2.53) is zero for $\omega = 2\pi k/N$.

Proof. The left side of (2.56) is zero for a periodic $w(\tau)$ at $\omega = 2\pi k/N$. ■

2.3 SIGNAL SPECTRA

The periodogram defines, in a sense, the frequency contents of a signal over a finite time interval. This information may, however, be fairly hidden due to the typically erratic behavior of a periodogram as a function of ω. We now seek a definition of a similar concept for signals over the interval $t \in [1,\infty)$. Preferably, such a concept should more clearly demonstrate the different frequency contributions to the signal.

A definition for our framework is, however, not immediate. It would perhaps be natural to define the spectrum of a signal s as

$$\lim_{N\to\infty} |S_N(\omega)|^2 \tag{2.57}$$

but this limit fails to exist for many signals of practical interest. Another possibility would be to use the concept of the spectrum, or spectral density, of a stationary stochastic process as the Fourier transform of its covariance function. However, the processes that we consider here are frequently not stationary, for reasons that are described later. We shall therefore develop a framework for describing signals and their spectra that is applicable to deterministic as well as stochastic signals.

A Common Framework for Deterministic and Stochastic Signals

In this book we shall frequently work with signals that are described as stochastic processes with deterministic components. The reason is, basically, that we prefer to consider the input sequence as deterministic, or at least partly deterministic, while disturbances on the system most conveniently are described by random variables. In this way the system output becomes a stochastic process with deterministic components. For (2.20) we find that

$$Ey(t) = G(q)u(t)$$

so $\{y(t)\}$ is not a stationary process.

To deal with this problem, we shall consider signals $\{s(t)\}$ subject to the following assumptions.

$$\text{(i)} \qquad Es(t) = m_s(t), \qquad |m_s(t)| \le C, \qquad \forall t$$

$$\text{(ii)} \qquad Es(t)s(r) = R_s(t, r), \qquad |R_s(t, r)| \le C \tag{2.58}$$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} R_s(t, t-\tau) = R_s(\tau), \qquad \forall \tau \tag{2.59}$$

When (i) and (ii) hold, we say that $\{s(t)\}$ is *quasi-stationary*. Here expectation E is with respect to the "stochastic components" of $s(t)$. If $\{s(t)\}$ itself is a deterministic sequence, the expectation is without effect and quasi-stationarity then means that $\{s(t)\}$ is a bounded sequence such that the limits

$$R_s(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} s(t)s(t-\tau)$$

exist. If $\{s(t)\}$ is a stationary stochastic process, (2.58) and (2.59) are trivially satisfied, since then $Es(t)s(t-\tau) \triangleq R_s(\tau)$ does not depend on t.

For easy notation we introduce the symbol $\bar{E}$ by

$$\bar{E}f(t) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} Ef(t) \tag{2.60}$$

with an implied assumption that the limit exists when the symbol is used. Assumption (2.59), which simultaneously is a definition of $R_s(\tau)$, then reads

$$\bar{E}s(t)s(t-\tau) = R_s(\tau) \tag{2.61}$$

Sometimes, with some abuse of notation, we shall call $R_s(\tau)$ the *covariance function* of s, keeping in mind that this is a correct term only if $\{s(t)\}$ is a stationary stochastic process with mean value zero.

Similarly, we say that two signals $\{s(t)\}$ and $\{w(t)\}$ are *jointly quasi-stationary* if they both are quasi-stationary and if, in addition, the *cross-covariance* function

$$R_{sw}(\tau) = \bar{E}s(t)w(t-\tau) \tag{2.62}$$

exists. We shall say that jointly quasi-stationary signals are *uncorrelated* if their cross-covariance function is identically zero.

Definition of Spectra

When limits like (2.61) and (2.62) hold, we define the (power) *spectrum* of $\{s(t)\}$ as

$$\Phi_s(\omega) = \sum_{\tau=-\infty}^{\infty} R_s(\tau)e^{-i\tau\omega} \tag{2.63}$$

and the *cross spectrum* between $\{s(t)\}$ and $\{w(t)\}$ as

$$\Phi_{sw}(\omega) = \sum_{\tau=-\infty}^{\infty} R_{sw}(\tau)e^{-i\tau\omega} \tag{2.64}$$

provided the infinite sums exist. In the sequel, as we talk of a signal's "spectrum," we always implicitly assume that the signal has all the properties involved in the definition of spectrum.

While $\Phi_s(\omega)$ always is real, $\Phi_{sw}(\omega)$ is in general a complex-valued function of ω. Its real part is known as the *cospectrum* and its imaginary part as the *quadrature spectrum*. The argument $\arg \Phi_{sw}(\omega)$ is called the *phase spectrum*, while $|\Phi_{sw}(\omega)|$ is the *amplitude spectrum*.

Note that, by definition of the inverse Fourier transform, we have

$$\bar{E}s^2(t) = R_s(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_s(\omega)d\omega \tag{2.65}$$

Example 2.3 Spectrum of a Sinusoid

Consider again the signal (2.45), now extended to the interval $[1,\infty)$. We have

$$\frac{1}{N}\sum_{k=1}^{N} Eu(k)u(k-\tau) = \frac{1}{N}\sum_{k=1}^{N} A^2\cos(\omega_0 k)\cos(\omega_0(k-\tau)), \qquad \tau \geq 0 \tag{2.66}$$

(Expectation is of no consequence since u is deterministic.) Now

$$\cos(\omega_0 k)\cos(\omega_0(k-\tau)) = \tfrac{1}{2}(\cos(2\omega_0 k - \omega_0\tau) + \cos\omega_0\tau)$$

which shows that

$$\bar{E}u(t)u(t-\tau) = \frac{A^2}{2}\cos\omega_0\tau = R_u(\tau)$$

The spectrum now is

$$\Phi_u(\omega) = \sum_{\tau=-\infty}^{\infty} \frac{A^2}{2} \cos(\omega_0 \tau) e^{-i\omega\tau} = \frac{A^2}{4}(\delta(\omega - \omega_0) + \delta(\omega + \omega_0)) \cdot 2\pi \quad (2.67)$$

where δ is Dirac's delta function. This result fits well with the finite interval expression (2.46). ■

Example 2.4 Stationary Stochastic Processes

Let $\{v(t)\}$ be a stationary stochastic process with covariance function (2.14). Since (2.59) then equals (2.14), our definition of spectrum coincides with the conventional one. Suppose now that the process v is given as (2.9). Its covariance function is then given by (2.13). The spectrum is

$$\begin{aligned}\Phi_v(\omega) &= \sum_{\tau=-\infty}^{\infty} \lambda e^{-i\tau\omega} \sum_{k=\max(0,\tau)}^{\infty} h(k)h(k-\tau) \\ &= \lambda \sum_{\tau=-\infty}^{\infty} \sum_{k=\max(0,\tau)}^{\infty} h(k)e^{-ik\omega} h(k-\tau)e^{i(k-\tau)\omega} \\ &= [k - \tau = s] = \lambda \sum_{s=0}^{\infty} h(s)e^{is\omega} \sum_{k=0}^{\infty} h(k)e^{-ik\omega} = \lambda |H(e^{i\omega})|^2\end{aligned}$$

using (2.18). This result is very important for our future use:

> The stochastic process described by $v(t) = H(q)e(t)$, where $\{e(t)\}$ is a sequence of independent random variables with zero mean values and covariances λ, has the spectrum
>
> $$\Phi_v(\omega) = \lambda |H(e^{i\omega})|^2 \quad (2.68)$$

This result, which was easy to prove for the special case of a stationary stochastic process, will be proved in the general case as Theorem 2.2 later in this section. Figure 2.7 shows the spectrum of the process of Figures 2.5 and 2.6, while the periodogram of the realization of Figure 2.6 is shown in Figure 2.8. ■

Example 2.5 Spectrum of a Mixed Deterministic and Stochastic Signal

Consider now a signal

$$s(t) = u(t) + v(t) \quad (2.69)$$

where $\{u(t)\}$ is a deterministic signal with spectrum $\Phi_u(\omega)$ and $\{v(t)\}$ is a stationary stochastic process with zero mean value and spectrum $\Phi_v(\omega)$. Then

$$\begin{aligned}\bar{E}s(t)s(t-\tau) &= \bar{E}u(t)u(t-\tau) + \bar{E}u(t)v(t-\tau) \\ &\quad + \bar{E}v(t)u(t-\tau) + \bar{E}v(t)v(t-\tau) = R_u(\tau) + R_v(\tau) \quad (2.70)\end{aligned}$$

since $\bar{E}v(t)u(t-\tau) = 0$. Hence

$$\Phi_s(\omega) = \Phi_u(\omega) + \Phi_v(\omega) \quad (2.71)$$

■

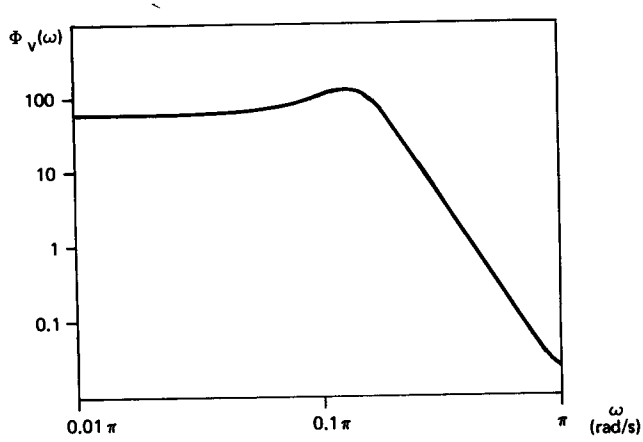

Figure 2.7 The spectrum of the process $v(t) - 1.5v(t-1) + 0.7v(t-2) = e(t) + 0.5e(t-1)$, $\{e(t)\}$ being white noise.

Connections to the Periodogram (*)

While the original idea (2.57) does not hold, a conceptually related result can be proved; that is, the expected value of the periodogram *converges weakly* to the spectrum:

$$E|S_N(\omega)|^2 \xrightarrow{w} \Phi_s(\omega) \tag{2.72}$$

By this is meant that

$$\lim_{N\to\infty} \int_{-\pi}^{\pi} E|S_N(\omega)|^2 \Psi(\omega)\, d\omega = \int_{-\pi}^{\pi} \Phi_s(\omega)\Psi(\omega)\, d\omega \tag{2.73}$$

for all sufficiently smooth functions $\Psi(\omega)$.

We have

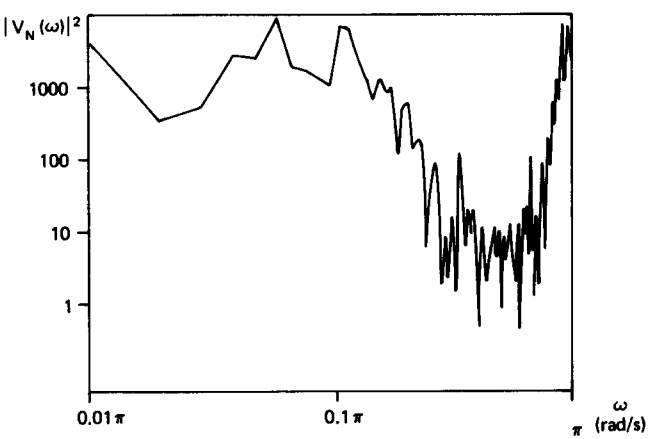

Figure 2.8 The periodogram of the realization of Figure 2.6.

Lemma 2.1. Suppose that $\{s(t)\}$ is quasi-stationary with spectrum $\Phi_s(\omega)$. Let

$$S_N(\omega) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} s(t)e^{-it\omega}$$

and let $\Psi(\omega)$ be an arbitrary function $|\omega| \leq \pi$ with Fourier coefficients a_τ such that

$$\sum_{\tau=-\infty}^{\infty} |a_\tau| < \infty$$

Then (2.73) holds.

Proof

$$\begin{aligned} E|S_N(\omega)|^2 &= \frac{1}{N} \sum_{k=1}^{N} \sum_{\ell=1}^{N} Es(k)s(\ell)e^{i\omega(k-\ell)} \\ &= [\ell - k = \tau] = \sum_{\tau=-(N-1)}^{N-1} R_N(\tau)e^{-i\omega\tau} \end{aligned} \tag{2.74}$$

where

$$R_N(\tau) = \frac{1}{N} \sum_{k=1}^{N} Es(k)s(k-\tau) \tag{2.75}$$

with the convention that $s(k)$ is taken as zero outside the interval $[1, N]$. Multiplying (2.74) by $\Psi(\omega)$ and integrating over $[-\pi, \pi]$ gives

$$\int_{-\pi}^{\pi} E|S_N(\omega)|^2\Psi(\omega)\, d\omega = \sum_{\tau=-(N-1)}^{N-1} R_N(\tau)a_\tau$$

by the definition of a_τ. Similarly, allowing interchange of summation and integration, we have

$$\int_{-\pi}^{\pi} \Phi_s(\omega)\Psi(\omega)\, d\omega = \sum_{\tau=-\infty}^{\infty} R_s(\tau)a_\tau$$

Hence

$$\begin{aligned} &\int_{-\pi}^{\pi} E|S_N(\omega)|^2\Psi(\omega)\, d\omega - \int_{-\pi}^{\pi} \Phi_s(\omega)\Psi(\omega)\, d\omega \\ &\qquad = \sum_{\tau=-(N-1)}^{N} a_\tau[R_N(\tau) - R_s(\tau)] + \sum_{|\tau|\geq N} a_\tau R_s(\tau) \end{aligned}$$

Problem 2D.5 now completes the proof. ■

Notice that for stationary stochastic processes the result (2.72) can be strengthened to "ordinary" convergence (see Problem 2D.3). Notice also that, in our framework, results like (2.72) can be applied also to realizations of stochastic processes simply by ignoring the expectation operator. We then view the realization in question as a given "deterministic" sequence, and will then, of course, have to require that the conditions (2.58) and (2.59) hold for this particular realization [disregard "E" also in (2.58) and (2.59)].

Transformation of Spectra by Linear Systems

As signals are filtered through linear systems, their properties will change. We saw how the periodogram was transformed in Theorem 2.1 and how white noise created stationary stochastic processes in (2.68). For spectra we have the following general result.

Theorem 2.2. Let $\{w(t)\}$ be a quasi-stationary signal with spectrum $\Phi_w(\omega)$, and let $G(q)$ be a stable transfer function. Let

$$s(t) = G(q)w(t) \tag{2.76}$$

Then $\{s(t)\}$ is also quasi-stationary and

$$\Phi_s(\omega) = |G(e^{i\omega})|^2 \Phi_w(\omega) \tag{2.77}$$

$$\Phi_{sw}(\omega) = G(e^{i\omega})\Phi_w(\omega) \tag{2.78}$$

Proof. The proof is given in Appendix 2A.

Corollary. Let $\{y(t)\}$ be given by

$$y(t) = G(q)u(t) + H(q)e(t) \tag{2.79}$$

where $\{u(t)\}$ is a quasi-stationary, deterministic signal with spectrum $\Phi_u(\omega)$, and $\{e(t)\}$ is white noise with variance λ. Let G and H be stable filters. Then $\{y(t)\}$ is quasi-stationary and

$$\Phi_y(\omega) = |G(e^{i\omega})|^2\Phi_u(\omega) + \lambda|H(e^{i\omega})|^2 \tag{2.80}$$

$$\Phi_{yu}(\omega) = G(e^{i\omega})\Phi_u(\omega) \tag{2.81}$$

Proof. The corollary follows from the theorem using Examples 2.4 and 2.5. ■

Spectral Factorization

Typically, the transfer functions $G(q)$ and $H(q)$ used here are rational functions of q. Then results like (2.68) and Theorem 2.2 describe spectra as real-valued rational functions of $e^{i\omega}$ (which means that they also are rational functions of $\cos\omega$).

In practice, the converse of such results is of major interest: Given a spectrum $\Phi_v(\omega)$, can we then find a transfer function $H(q)$ such that the process $v(t) = H(q)e(t)$ has this spectrum with $\{e(t)\}$ being white noise? It is quite clear that this is not possible for all positive functions $\Phi_v(\omega)$. For example, if the spectrum is zero on an interval, then the function $H(z)$ must be zero on a portion of the unit circle. But since by necessity $H(z)$ should be analytic outside and on the unit circle for the expansion (2.18) to make sense, this implies that $H(z)$ is zero everywhere and cannot match the chosen spectrum.

The exact conditions under which our question has a positive answer are

discussed in texts on stationary processes, such as Wiener (1949) and Rozanov (1967). For our purposes it is sufficient to quote a simpler result, dealing only with spectral densities $\Phi_v(\omega)$ that are rational in the variable $e^{i\omega}$ (or $\cos\omega$).

Spectral factorization: Suppose that $\Phi_v(\omega) > 0$ is a rational function of $\cos\omega$ (or $e^{i\omega}$). Then there exists a monic rational function of z, $R(z)$, with no poles and no zeros on or outside the unit circle such that

$$\Phi_v(\omega) = \lambda |R(e^{i\omega})|^2$$

The proof of this result consists of a straightforward construction of R, and it can be found in standard texts on stochastic processes or stochastic control (e.g., Rozanov, 1967; Åström, 1970).

Example 2.6 ARMA Processes

If a stationary process $\{v(t)\}$ has rational spectrum $\Phi_v(\omega)$, we can thus represent it as

$$v(t) = R(q)e(t) \tag{2.82}$$

where $\{e(t)\}$ is white noise with variance λ. Here $R(q)$ is a rational function

$$R(q) = \frac{C(q)}{A(q)}$$

$$C(q) = 1 + c_1 q^{-1} + \cdots + c_{n_c} q^{-n_c}$$

$$A(q) = 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a}$$

so that we may write

$$v(t) + a_1 v(t-1) + \cdots + a_{n_a} v(t-n_a) = e(t) + c_1 e(t-1) + \cdots + c_{n_c} e(t-n_c) \tag{2.83}$$

for (2.82). Such a representation of a stochastic process is known as an *ARMA model.* If $n_c = 0$, we have an autoregressive (AR) model:

$$v(t) + a_1 v(t-1) + \cdots + a_{n_a} v(t-n_a) = e(t) \tag{2.84}$$

And if $n_a = 0$, we have a moving average (MA) model:

$$v(t) = e(t) + c_1 e(t-1) + \cdots + c_{n_c} e(t-n_c) \tag{2.85}$$

■

The spectral factorization concept is important since it provides a way of representing the disturbance in the standard form $v = H(q)e$ from information about its spectrum only. The spectrum is usually a sound engineering way of describing properties of signals: "The disturbances are concentrated around 50 Hz" or "We are having low-frequency disturbances with little effect over 1 rad/s." Rational functions are able to approximate functions of rather versatile shapes. Hence the spectral factorization result will provide a good modeling framework for disturbances.

Second-order Properties

The signal spectra, as defined here, describe the *second-order properties* of the signals (for stochastic processes, their second-order statistics, i.e., first and second moments). Recall from Section 2.1 that stochastic processes may have very different looking realizations even if they have the same covariance function (see Figures 2.5 and 2.6)! The spectrum thus describes only certain aspects of a signal. Nevertheless, it will turn out that many properties related to identification depend only on the spectra of the involved signals. This motivates our detailed interest in the second-order properties.

2.4 SINGLE REALIZATION BEHAVIOR AND ERGODICITY RESULTS (*)

All the results of the previous section are also valid, as we pointed out, for the special case of a given deterministic signal $\{s(t)\}$. Definitions of spectra, their transformations (Theorem 2.2) and their relationship with the periodogram (Lemma 2.1) hold unchanged; we may just disregard the expectation E and interpret $\bar{E}f(t)$ as

$$\lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} f(t)$$

There is a certain charm with results like these that do not rely on a probabilistic framework: we anyway observe just one realization, so why should we embed this observation in a stochastic process and describe its average properties taken over an ensemble of potential observations? There are two answers to this question. One is that such a framework facilitates certain calculations. Another is that it allows us to deal with the question of what would happen if the experiment were repeated.

Nevertheless, it is a valid question to ask whether the spectrum of the signal $\{s(t)\}$, as defined in a probabilistic framework, differs from the spectrum of the actually observed, single realization were it to be considered as a given, deterministic signal. This is the problem of *ergodic* theory, and for our setup we have the following fairly general result.

Theorem 2.3. Let $\{s(t)\}$ be a quasi-stationary signal. Let $Es(t) = m(t)$. Assume that

$$s(t) - m(t) = v(t) = \sum_{k=0}^{\infty} h_t(k)e(t-k) = H_t(q)e(t) \tag{2.86}$$

where $\{e(t)\}$ is a sequence of independent random variables with zero mean values, $Ee^2(t) = \lambda_t$, and bounded fourth moments, and where $\{H_t(q), t = 1,2,\dots\}$ is a uniformly stable family of filters. Then, with probability 1 as N tends to infinity,

$$\frac{1}{N} \sum_{t=1}^{N} s(t)s(t-\tau) \to \bar{E}\, s(t)s(t-\tau) = R_s(\tau) \tag{2.87a}$$

$$\frac{1}{N}\sum_{t=1}^{N}[s(t)m(t-\tau)-Es(t)m(t-\tau)]\to 0 \tag{2.87b}$$

$$\frac{1}{N}\sum_{t=1}^{N}[s(t)v(t-\tau)-Es(t)v(t-\tau)]\to 0 \tag{2.87c}$$

The proof is given in Appendix 2B. ■

The theorem is quite important. It says that, provided the stochastic part of the signal can be described as filtered white noise as in (2.86), then

> the spectrum of an observed single realization of $\{s(t)\}$, computed as for a deterministic signal, coincides, with probability 1, with that of the process $\{s(t)\}$, defined by ensemble averages (E) as in (2.61).

This de-emphasizes the distinction between deterministic and stochastic signals when we consider second-order properties only. A signal $\{s(t)\}$ whose spectrum is $\Phi_s(\omega) \equiv \lambda$ may, for all purposes related to second-order properties, be regarded as a realization of white noise with variance λ.

The theorem also gives an answer to the question of whether our "theoretical" spectrum, defined in (2.63) using the physically unrealizable concepts of E and lim, relates to the actually observed periodogram (2.43). According to Theorem 2.3 and Lemma 2.1, "smoothed" versions of $|S_N(\omega)|^2$ will look like $\Phi_s(\omega)$ for large N. Compare Figures 2.7 and 2.8. This link between our theoretical concepts and the real data is of course of fundamental importance. See Section 6.3.

2.5 MULTIVARIABLE SYSTEMS (*)

So far, we have worked with systems having a scalar input and a scalar output. In this section we shall consider the case where the output signal has p components and the input signal has m components. Such systems are called *multivariable*. The extra work involved in dealing with models of multivariable systems can be split up into two parts:

1. The easy part: mostly notational changes, keeping track of transposes, and noting that certain scalars become matrices and might not commute.
2. The difficult part: multioutput models have a much richer internal structure, which has the consequence that their parametrization is nontrivial. See Appendix 4A. (Multiple-input, single-output, MISO, models do not expose these problems.)

Let us collect the p components of the output signal into a p-dimensional column vector $y(t)$ and similarly construct an m-dimensional input vector $u(t)$. Let the disturbance $e(t)$ also be a p-dimensional column vector. The basic system description then looks just like (2.20):

$$y(t) = G(q)u(t) + H(q)e(t) \tag{2.88}$$

where now $G(q)$ is a transfer function matrix of dimension $p \times m$ and $H(q)$ has dimension $p \times p$. This means that the i, j entry of $G(q)$, denoted by $G_{ij}(q)$, is the scalar transfer function from input number j to output number i. The sequence $\{e(t)\}$ is a sequence of independent random p-dimensional vectors with zero mean values and covariance matrices $E\, e(t)e^T(t) = \Lambda$.

Now, all the development in this chapter goes through with proper interpretation of matrix dimensions. Note in particular the following:

- The impulse responses $g(k)$ and $h(k)$ will be $p \times m$ and $p \times p$ matrices, respectively, with norms

$$\|g(k)\| = \left(\sum_{i,j} |g_{ij}|^2\right)^{1/2} \tag{2.89}$$

  replacing absolute values in the definitions of stability.

- The definitions of covariances become [cf. (2.59)]

$$\bar{E}\, s(t)s^T(t - \tau) = R_s(\tau) \tag{2.90}$$

$$\bar{E}\, s(t)w^T(t - \tau) = R_{sw}(\tau) \tag{2.91}$$

  These are now matrices, with norms as in (2.89).

- Definitions of spectra remain unchanged, while the counterpart of the corollary to Theorem 2.2 reads

$$\Phi_y(\omega) = G(e^{i\omega})\Phi_u(\omega)G^T(e^{-i\omega}) + H(e^{i\omega})\Lambda H^T(e^{-i\omega}) \tag{2.92}$$

  Notice that the definition of spectra for vector signals implicitly defines also cross-spectra between the components of the signal. See also Problem 2G.3.

- The spectral factorization result now reads: Suppose that $\Phi_v(\omega)$ is a $p \times p$ matrix that is positive definite for all ω and whose entries are rational functions of $\cos \omega$ (or $e^{i\omega}$). Then there exists a $p \times p$ monic matrix function $H(z)$ whose entries are rational functions of z (or z^{-1}) such that the (rational) function $\det H(z)$ has no poles and no zeros on or outside the unit circle. (For a proof, see Theorem 10.1 in Rozanov, 1967).

- The formulation of Theorem 2.3 carries over without changes. (In fact, the proof in Appendix 2B is carried out for the multivariable case).

2.6 SUMMARY

We have established the representation

$$y(t) = G(q)u(t) + H(q)e(t) \tag{2.93}$$

as the basic description of a linear system subject to additive random disturbances. Here $\{e(t)\}$ is a sequence of independent random variables with zero mean values and variances λ (in the multioutput case, covariance matrices Λ). Also,

$$G(q) = \sum_{k=1}^{\infty} g(k)q^{-k}$$

$$H(q) = 1 + \sum_{k=1}^{\infty} h(k)q^{-k}$$

The filter $G(q)$ is *stable* if

$$\sum_{k=1}^{\infty} |g(k)| < \infty$$

As the reader no doubt is aware of, other particular ways of representing linear systems, such as state-space models and difference equations, are quite common in practice. These can, however, be viewed as particular ways of representing the sequences $\{g(k)\}$ and $\{h(k)\}$, and they will be dealt with in some detail in Chapter 4.

We have also discussed the frequency function $G(e^{i\omega})$, bearing information about how an input sinusoid of frequency ω is transformed by the system. Frequency-domain concepts in terms of the frequency contents of input and output signals were also treated. The Fourier transform of a finite-interval signal was defined as

$$U_N(\omega) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} u(t)e^{-i\omega t} \tag{2.94}$$

A signal $s(t)$ such that the limits

$$\bar{E}s(t)s(t-\tau) = R_s(\tau)$$

exist was said to be *quasi-stationary.*

$$(\bar{E}f(t) = \lim_{N\to\infty} \frac{1}{N} \sum_{1}^{N} Ef(t))$$

Then the *spectrum* of $s(t)$ is defined as

$$\Phi_s(\omega) = \sum_{\tau=-\infty}^{\infty} R_s(\tau)e^{-i\omega\tau} \tag{2.95}$$

For y generated as in (2.93) with $\{u(t)\}$ and $\{e(t)\}$ independent, we then have

$$\Phi_y(\omega) = |G(e^{i\omega})|^2\Phi_u(\omega) + \lambda|H(e^{i\omega})|^2$$

2.7 BIBLIOGRAPHY

The material of this chapter is covered in many textbooks on systems and signals. For a thorough elementary treatment, see Oppenheim and Willsky (1983). A discussion oriented more toward signals as time series is given in Brillinger (1981), which also contains several results of the same character as our Theorems 2.1 and 2.2.

A detailed discussion of the sampling procedure and connections between the physical time-continuous system and the sampled-data description (2.6) is given in Chapter 4 of Åström and Wittenmark (1984). Chapter 6 of that book also contains an illuminating discussion of disturbances and how to describe them mathematically. The idea of describing stochastic disturbances as linearly filtered white noise goes back to Wold (1938).

Fourier techniques in the analysis and description of signals go back to the Babylonians. See Oppenheim and Willsky (1983), Section 4.0, for a brief historical account. The periodogram was evidently introduced by Schuster (1894) to study periodic phenomena without having to consider relative phases. The statistical properties of the periodogram were first studied by Slutsky (1929). See also Brillinger (1983). Concepts of spectra are intimately connected to the harmonic analysis of time series, as developed by Wiener (1930), Wold (1938), Kolmogorov (1941a), and others. Useful textbooks on these concepts (and their estimation) include Jenkins and Watts (1968) and Brillinger (1981). Our definition of the Fourier transform (2.37) with summation from 1 to N and a normalization with $1/\sqrt{N}$ suits our purposes, but is not standard. The placement of 2π in the definition of the spectrum or in the inverse transform, as we have it in (2.65), varies in the literature. Our choice is based on the wish to let white noise have a constant spectrum whose value equals the variance of the noise. The particular framework chosen here to accommodate mixtures of stochastic processes and deterministic signals is apparently novel, but has a close relationship to the classical concepts.

The result of Theorem 2.2 is standard when applied to stationary stochastic processes. See, for example, James, Nichols, and Phillips (1947) or Åström (1970). The extension to quasi-stationary signals appears to be new.

Spectral factorization turned out to be a key issue in the prediction of time series. It was formulated and solved by Wiener (1949) and Paley and Wiener (1934). The multivariable version is treated in Youla (1961). The concept is now standard in textbooks on stationary processes (see, e.g., Rozanov, 1967).

The topic of single realization behavior is a standard problem in probability theory. See, for example, Ibragimov and Linnik (1971), Billingsley (1965), or Chung (1974) for general treatments of such problems.

2.8 PROBLEMS†

2G.1. Let $s(t)$ be a p-dimensional signal. Show that

$$\overline{E}|s(t)|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} \operatorname{tr}(\Phi_s(\omega))\, d\omega$$

2G.2. Let $\Phi_s(\omega)$ be the (power) spectrum of a scalar signal defined as in (2.63). Show that

(i) $\Phi_s(\omega)$ is real.

(ii) $\Phi_s(\omega) \geq 0 \; \forall \omega$.

(iii) $\Phi_s(-\omega) = \Phi_s(\omega)$.

† See the Preface for an explanation of the numbering system.

2G.3. Let $s(t) = \begin{bmatrix} y(t) \\ u(t) \end{bmatrix}$ and let its spectrum be

$$\Phi_s(\omega) = \begin{bmatrix} \Phi_y(\omega) & \Phi_{yu}(\omega) \\ \Phi_{uy}(\omega) & \Phi_u(\omega) \end{bmatrix}$$

Show that $\Phi_s(\omega)$ is a Hermitian matrix: that is,

$$\Phi_s(\omega) = \Phi_s^*(\omega)$$

where * denotes transpose and complex conjugate. What does this imply about the relationships between the cross spectra $\Phi_{yu}(\omega)$, $\Phi_{uy}(\omega)$, and $\Phi_{yu}(-\omega)$?

2G.4. Let a continuous time system representation be given by

$$y(t) = G_c(p)u(t)$$

The input is constant over the sampling interval T. Show that the sampled input-output data are related by

$$y(t) = G_T(q)u(t)$$

where

$$G_T(q) = \int_{s=-i\infty}^{i\infty} G_c(s) \frac{e^{sT} - 1}{s} \frac{1}{q - e^{sT}} ds$$

Hint: Use (2.5).

2E.1. A stationary stochastic process has the spectrum

$$\Phi_v(\omega) = \frac{1.25 + \cos\omega}{1.64 + 1.6\cos\omega}$$

Describe $\{v(t)\}$ as an ARMA process.

2E.2. Suppose that $\{\eta(t)\}$ and $\{\xi(t)\}$ are two mutually independent sequences of independent random variables with

$$E\eta(t) = E\xi(t) = 0, \qquad E\eta^2(t) = \lambda_\eta, \qquad E\xi^2(t) = \lambda_\xi$$

Consider

$$w(t) = \eta(t) + \xi(t) + \gamma\xi(t-1)$$

Determine a MA(1) process

$$v(t) = e(t) + ce(t-1)$$

where $\{e(t)\}$ is white noise with

$$Ee(t) = 0, \qquad Ee^2(t) = \lambda_e$$

such that $\{w(t)\}$ and $\{v(t)\}$ have the same spectra; that is, find c and λ_e so that $\Phi_v(\omega) \equiv \Phi_w(\omega)$.

2E.3. (a) In Problem 2E.2 assume that $\{\eta(t)\}$ and $\{\xi(t)\}$ are jointly Gaussian. Show that if $\{e(t)\}$ also is chosen as Gaussian then the joint probability distribution of the process $\{w(t)\}$ [i.e., the joint PDFs of $w(t_1), w(t_2), \ldots, w(t_p)$ for any collection of

time instances t_i] coincides with that of the process $\{v(t)\}$. Then, for all practical purposes, *the processes $\{v(t)\}$ and $\{w(t)\}$ are indistinguishable.*

(b) Assume now that $\eta(t) \in N(0,\lambda_\eta)$, while

$$\xi(t) = \begin{cases} 1, & \text{w.p.} \quad \dfrac{\lambda_\xi}{2} \\ -1, & \text{w.p.} \quad \dfrac{\lambda_\xi}{2} \\ 0, & \text{w.p.} \quad 1 - \lambda_\xi \end{cases}$$

Show that, although v and w have the same spectra, we cannot find a distribution for $e(t)$ so that they have the same joint PDFs. *Consequently the process $w(t)$ cannot be represented as an MA(1) process, although it has a second-order equivalent representation of that form.*

2E.4. Consider the "state-space description"

$$\begin{aligned} x(t+1) &= fx(t) + w(t) \\ y(t) &= hx(t) + v(t) \end{aligned}$$

where x, f, h, w, and v are scalars. $\{w(t)\}$ and $\{v(t)\}$ are mutually independent white Gaussian noises with variances R_1 and R_2, respectively. Show that $y(t)$ can be represented as an ARMA process:

$$y(t) + a_1 y(t-1) + \cdots + a_n y(t-n) = e(t) + c_1 e(t-1) + \cdots + c_n e(t-n)$$

Determine n, a_i, c_i, and the variance of $e(t)$ in terms of f, h, R_1, and R_2. What is the relationship between $e(t)$, $w(t)$, and $v(t)$?

2E.5. Consider the system

$$y(t) = G(q)u(t) + v(t)$$

controlled by the regulator

$$u(t) = -F_2(q)y(t) + F_1(q)r(t)$$

where $\{r(t)\}$ is a quasi-stationary reference signal with spectrum $\Phi_r(\omega)$. The disturbance $\{v(t)\}$ has spectrum $\Phi_v(\omega)$. Assume that $\{r(t)\}$ and $\{v(t)\}$ are uncorrelated and that the resulting closed-loop system is stable. Determine the spectra $\Phi_y(\omega)$, $\Phi_u(\omega)$, and $\Phi_{yu}(\omega)$.

2E.6. Consider the system

$$\frac{d}{dt}y(t) + ay(t) = u(t) \tag{2.96}$$

Suppose that the input $u(t)$ is piecewise constant over the sampling interval

$$u(t) = u_k, \qquad kT \le t < (k+1)T$$

(a) Derive a sampled-data system description for u_k, $y(kT)$.

(b) Assume that there is a time delay of T seconds so that $u(t)$ in (2.96) is replaced by $u(t-T)$. Derive a sampled-data system description for this case.

(c) Assume that the time delay is $1.5T$ so that $u(t)$ is replaced by $u(t - 1.5T)$. Then give the sampled-data description.

2E.7. Consider a system given by

$$y(t) + ay(t-1) = bu(t-1) + e(t) + ce(t-1)$$

where $\{u(t)\}$ and $\{e(t)\}$ are independent white noises, with variances μ and λ, respectively. Follow the procedure suggested in Appendix 2C to multiply the system description by $e(t)$, $e(t-1)$, $u(t)$, $u(t-1)$, $y(t)$, and $y(t-1)$, respectively, and take expectation to show that

$$R_{ye}(0) = \lambda, \qquad R_{ye}(1) = (c - a)\lambda$$

$$R_{yu}(0) = 0, \qquad R_{yu}(1) = b\mu$$

$$R_y(0) = \frac{b^2\mu + \lambda + c^2\lambda - 2ac\lambda}{1 - a^2}, \qquad R_y(1) = \frac{\lambda(c - a + a^2c - ac^2) - ab^2\mu}{1 - a^2}$$

2T.1. Consider a continuous time system (2.1):

$$y(t) = \int_{\tau=0}^{\infty} g(\tau)u(t-\tau)\,d\tau$$

Let $g_T(\ell)$ be defined by (2.5), and assume that $u(t)$ is not piecewise constant, but that

$$\left|\frac{d}{dt}u(t)\right| \le C_1$$

Let $u_k = u\left((k + \frac{1}{2})T\right)$

and show that

$$y(kT) = \sum_{\ell=1}^{\infty} g_T(\ell)u_{k-\ell} + r_k$$

where

$$|r_k| \le C_2 \cdot T^2$$

Give a bound for C_2.

2T.2. If the filters $R_1(q)$ and $R_2(q)$ are (strictly) stable, then show that $R_1(q)R_2(q)$ is also (strictly) stable (see also Problem 3D.1).

2T.3. Let $G(q)$ be a rational transfer function; that is,

$$G(q) = \frac{b_1q^{n-1} + \cdots + b_n}{q^n + a_1q^{n-1} + \cdots + a_n}$$

Show that if $G(q)$ is stable, then it is also strictly stable.

2T.4. Consider the time-varying system

$$x(t+1) = F(t)x(t) + G(t)u(t)$$
$$y(t) = H(t)x(t)$$

Write

$$y(t) = \sum_{k=1}^{t} g_t(k)u(t-k)$$

[take $g_t(k) = 0$ for $k > t$]. Assume that

$$F(t) \to \bar{F}, \qquad \text{as } t \to \infty$$

where $\bar{F}$ has all eigenvalues inside the unit circle. Show that the family of filters $\{g_t(k), t = 1, 2, \ldots\}$ is uniformly stable.

2D.1. Consider $U_N(\omega)$ defined by (2.37). Show that $U_N(2\pi - \omega) = U_N(\omega) = \bar{U}_N(\omega)$ and rewrite (2.38) in terms of real-valued quantities only.

2D.2. Establish (2.39).

2D.3. Let $\{u(t)\}$ be a stationary stochastic process with $R_u(\tau) = Eu(t)u(t-\tau)$, and let $\Phi_u(\omega)$ be its spectrum. Assume that

$$\sum_{1}^{\infty} |\tau R_u(\tau)| < \infty$$

Let $U_N(\omega)$ be defined by (2.37). Prove that

$$E|U_N(\omega)|^2 \to \Phi_u(\omega), \qquad \text{as } N \to \infty$$

This is a strengthening of Lemma 2.1 for stationary processes.

2D.4. Let $G(q)$ be a stable system. Prove that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} k|g(k)| = 0$$

Hint: Use Kronecker's lemma: Let a_k, b_k be sequences such that a_k is positive and decreasing to zero. Then $\Sigma a_k b_k < \infty$ implies

$$\lim_{N\to\infty} a_N \sum_{1}^{N} b_k = 0$$

(see, e.g., Chung, 1974, for a proof of Kronecker's lemma).

2D.5. Let $b_N(\tau)$ be a doubly indexed sequence such that, $\forall\tau$,

$$b_N(\tau) \to b(\tau), \qquad \text{as } N \to \infty$$

(but not necessarily uniformly in τ). Let a_τ be an infinite sequence, and assume that

$$\sum_{1}^{\infty} |a_\tau| < \infty, \qquad |b(\tau)| \le C \qquad \forall\tau$$

Show that

$$\lim_{N\to\infty} \left[\sum_{\tau=-N}^{N} a_\tau (b_N(\tau) - b(\tau)) + \sum_{|\tau|>N} a_\tau b(\tau) \right] = 0$$

Hint: Study Appendix 2A.

2S.1. In a series of problems in this book we shall develop the basis for an interactive

software package for system identification. The problems in question are marked with the letter *S*. For reasonable effort, it is assumed that the programming is performed in a high-level environment, such as APL or PC-MATLAB, with basic plotting and matrix handling routines available.

Write a MACRO

BODEPLOT(G)

that plots the Bode diagram of the transfer function G, entered as a complex-valued vector $G(e^{i\omega_k})$, $k = 1, 2\ldots, N$. Include the option of plotting several curves in the same diagram.

2C.1. Use the ideas of Appendix 2C to write a program that computes covariances for an arbitrary ARMA process based on algorithms for linear equations.

APPENDIX 2A: PROOF OF THEOREM 2.2

We carry out the proof for the multivariate case. Let $w(s) = 0$ for $s \leq 0$, and consider

$$R_s^N(\tau) = \frac{1}{N}\sum_{t=1}^{N} Es(t)s^T(t-\tau) = \frac{1}{N}\sum_{t=1}^{N}\sum_{k=0}^{t}\sum_{\ell=0}^{t-\tau} g(k)Ew(t-k)w^T(t-\tau-\ell)g^T(\ell) \tag{2A.1}$$

With the convention that $w(s) = 0$ if $s \notin [0, N]$, we can write

$$R_s^N(\tau) = \sum_{k=0}^{N}\sum_{\ell=0}^{N} g(k)\frac{1}{N}\sum_{t=1}^{N} Ew(t-k)w^T(t-\tau-\ell)g^T(\ell) \tag{2A.2}$$

If $w(s) \neq 0$, $s \leq 0$, then $s(t)$ gets the negligible contribution

$$\tilde{s}(t) = \sum_{k=t}^{\infty} g(k)w(t-k)$$

Let

$$R_w^N(\tau) = \frac{1}{N}\sum_{t=1}^{N} Ew(t)w^T(t-\tau)$$

We see that $R_w^N(\tau + \ell - k)$ and the inner sum in (2A.2) differ by at most $\max(k, |\tau + \ell|)$ summands, each of which are bounded by C according to (2.58). Thus

$$|R_w^N(\tau+\ell-k) - \frac{1}{N}\sum_{t=1}^{N} Ew(t-k)w^T(t-\tau-\ell)| \leq C\frac{\max(k, |\tau+\ell|)}{N} \leq \frac{C}{N}(k + |\tau+\ell|) \tag{2A.3}$$

Let us define

$$R_s(\tau) = \sum_{k=0}^{\infty}\sum_{\ell=0}^{\infty} g(k)R_w(\tau+\ell-k)g^T(\ell) \tag{2A.4}$$

Then

$$R_s(\tau) - R_s^N(\tau) \leq \sum_{k=N+1}^{\infty} \sum_{\ell=N+1}^{\infty} |g(k)||g(\ell)||R_w(\tau+\ell-k)|$$
$$+ \sum_{k=0}^{N} \sum_{\ell=0}^{N} |g(k)||g(\ell)||R_w(\tau+\ell-k) - R_w^N(\tau+\ell-k)|$$
$$+ \frac{C}{N} \sum_{k=0}^{N} k|g(k)| \cdot \sum_{\ell=0}^{N} |g(\ell)|$$
$$+ \frac{C}{N} \sum_{\ell=0}^{N} |\tau+\ell||g(\ell)| \cdot \sum_{k=0}^{N} |g(k)| \qquad (2A.5)$$

The first sum tends to zero as $N \to \infty$ since $|R_w(\tau)| \leq C$ and $G(q)$ is stable. It follows from the stability of $G(q)$ that

$$\frac{1}{N} \sum_{k=0}^{N} k|g(k)| \to 0, \qquad \text{as } N \to \infty \qquad (2A.6)$$

(see Problem 2D.4). Hence the last two sums of (2A.5) tend to zero as $N \to \infty$. Consider now the second sum of (2A.5). Select an arbitrary $\varepsilon > 0$, and choose $N = N_\varepsilon$ such that

$$\sum_{k=N_\varepsilon+1}^{\infty} |g(k)| < \frac{\varepsilon}{[C \cdot C_1]} \qquad (2A.7)$$

where

$$C_1 = \sum_{k=0}^{\infty} |g(k)|$$

This is possible since G is stable. Then select N'_ε such that

$$\max_{\substack{1 \leq \ell \leq N_\varepsilon \\ 1 \leq k \leq N_\varepsilon}} |R_w(\tau+\ell-k) - R_w^N(\tau+\ell-k)| < \frac{\varepsilon}{C_1^2}$$

for $N > N'_\varepsilon$. This is possible since

$$R_w^N(\tau) \to R_w(\tau), \qquad \text{as } N \to \infty \qquad (2A.8)$$

(w is quasi-stationary) and since only a finite number of $R_w(s)$'s are involved [no uniform convergence of (2A.8) is necessary]. Then, for $N > N'_\varepsilon$, we have that the second sum of (2A.5) is bounded by

$$\sum_{k=0}^{N_\varepsilon} \sum_{\ell=0}^{N_\varepsilon} |g(k)||g(\ell)| \cdot \frac{\varepsilon}{C_1^2} + \sum_{k=N_\varepsilon+1}^{\infty} \sum_{\ell=0}^{\infty} |g(k)||g(\ell)| \cdot 2C + \sum_{k=0}^{N_\varepsilon} \sum_{\ell=N_\varepsilon+1}^{\infty} |g(k)||g(\ell)| \cdot 2C$$

which is less than 5ε according to (2A.7). Hence, also, the second sum of (2A.5) tends to zero as $N \to \infty$, and we have proved that the limit of (2A.5) is zero and hence that $s(t)$ is quasi-stationary.

The proof that $Es(t)w^T(t-\tau)$ exists is analogous and simpler. For $\Phi_s(\omega)$ we now find that

$$\begin{aligned}\Phi_s(\omega) &= \sum_{\tau=-\infty}^{\infty}\Big(\sum_{k=0}^{\infty}\sum_{\ell=0}^{\infty} g(k)R_w(\tau+\ell-k)g^T(\ell)\Big)e^{-i\tau\omega} \\ &= \sum_{\tau=-\infty}^{\infty}\sum_{k=0}^{\infty} g(k)e^{-ik\omega}\sum_{\ell=0}^{\infty} R_w(\tau-\ell+k)e^{-i(\tau+\ell-k)\omega}g^T(\ell)e^{i\ell\omega} \\ &= [\tau-\ell+k=s] = \sum_{k=0}^{\infty} g(k)e^{-ik\omega}\cdot\sum_{s=-\infty}^{\infty} R_w(s)e^{-is\omega}\cdot\sum_{\ell=0}^{\infty} g^T(\ell)e^{i\ell\omega} \\ &= G(e^{i\omega})\Phi_w(\omega)G^T(e^{-i\omega})\end{aligned}$$

Hence (2.77) is proved. The result (2.78) is analogous and simpler.

APPENDIX 2B: PROOF OF THEOREM 2.3

In this appendix we shall show a more general variant of Theorem 2.3, which will be of value for the convergence analysis of Chapter 8. We also treat the multivariable case.

Theorem 2B.1. Let $\{G_\theta(q),\ \theta\in D_\theta\}$ be a uniformly stable family of filters, and assume that the family of deterministic signals $\{w_\theta(t)\}$, $\theta\in D_\theta$, $t=1,2,\dots,$ is subject to

$$|w_\theta(t)|\le C_w, \qquad \forall\theta,\ \forall t \tag{2B.1}$$

Let the signal $s_\theta(t)$ be defined, for each $\theta\in D_\theta$, by

$$s_\theta(t)=G_\theta(q)v(t)+w_\theta(t) \tag{2B.2}$$

where $\{v(t)\}$ is subject to the conditions of Theorem 2.3 [see (2.86) and let $Ee(t)e^T(t)=\Lambda_t$]. Then

$$\sup_{\theta\in D_\theta}\left\|\frac{1}{N}\sum_{t=1}^{N}[s_\theta(t)s_\theta^T(t)-Es_\theta(t)s_\theta^T(t)]\right\|\to 0 \quad \text{w.p. 1}, \qquad \text{as } N\to\infty \tag{2B.3}$$

Remark. We note that with $\dim s=1$, $D_\theta=\{\theta^*\}$ (only one element), $G_{\theta^*}(q)=1$, and $w_{\theta^*}(t)=m(t)$, then (2B.3) implies (2.87a). With

$$s_{\theta^*}(t)=\begin{bmatrix}s(t)\\ m(t)\\ v(t)\end{bmatrix}=\begin{bmatrix}1\\0\\1\end{bmatrix}v(t)+\begin{bmatrix}w_{\theta^*}(t)\\ w_{\theta^*}(t)\\ 0\end{bmatrix} \tag{2B.4}$$

the different cross products in (2B.3) imply all the results (2.87). ■

To prove Theorem 2B.1, we first establish two lemmas.

Lemma 2B.1. Let $\{v(t)\}$ obey the conditions of Theorem 2.3 and let

$$C_H = \sum_{k=1}^{\infty} \sup_t |h_t(k)|, \qquad C_e = \sup_t E|e(t)|^4, \qquad C_w = \sup_{t,\theta} |w_\theta(t)|$$

Then, for all r, N, m, and n,

$$E\left\| \sum_{t=r}^{N} [v(t-m)v^T(t-n) - Ev(t-m)v^T(t-n)] \right\|^2 \le C \cdot C_H^4 \cdot (N-r) \tag{2B.5}$$

$$E\left\| \sum_{t=r}^{N} v(t-m)w_\theta^T(t-n) \right\|^2 \le C \cdot C_H^2 \cdot C_w^2 \cdot (N-r) \tag{2B.6}$$

Proof of Lemma 2B.1. With no loss of generality, we may take $m = n = 0$. We then have

$$S_r^N \triangleq \sum_{t=r}^{N} v(t)v^T(t) - Ev(t)v^T(t) = \sum_{t=r}^{N} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} h_t(k)\alpha(t,k,\ell)h_t^T(\ell) \tag{2B.7}$$

where

$$\alpha(t,k,\ell) = e(t-k)e^T(t-\ell) - \Lambda_{t-\ell}\delta_{k\ell} \tag{2B.8}$$

For the square of the i,j entry of the matrix (2B.7), we have

$$(S_r^N(i,j))^2 = \sum_{t=r}^{N} \sum_{s=r}^{N} \sum_{k_1=0}^{\infty} \sum_{\ell_1=0}^{\infty} \sum_{k_2=0}^{\infty} \sum_{\ell_2=0}^{\infty} \gamma(t,s,k_1,k_2,\ell_1,\ell_2)$$

with

$$\gamma(t,s,k_1,k_2,\ell_1,\ell_2) = h_t^{(i)}(k_1)\alpha(t,k_1,\ell_1)[h_t^{(j)}(\ell_1)]^T h_s^{(i)}(k_2)\alpha(s,k_2,\ell_2)[h_s^{(j)}(\ell_2)]^T$$

Superscript (i) indicates the ith row vector. Since $\{e(t)\}$ is a sequence of independent variables, the expectation of γ is zero, unless at least some of the time indexes involved in $\alpha(t,k_1,\ell_1)$ and $\alpha(s,k_2,\ell_2)$ coincide, that is, unless

$$t-k_1 = s-k_2 \quad \text{or} \quad t-k_1 = s-\ell_2 \quad \text{or} \quad t-\ell_1 = s-k_2 \quad \text{or} \quad t-\ell_1 = s-\ell_2$$

For given values of t, k_1, k_2, ℓ_1, and ℓ_2, this may happen for at most four values of s. For these we also have

$$E\gamma(t,s,k_1,k_2,\ell_1,\ell_2) \le C_e \cdot |h(k_1)| \cdot |h(k_2)| \cdot |h(\ell_1)| \cdot |h(\ell_2)|$$

Hence

$$E(S_r^N(i,j))^2 \le \sum_{k_1=0}^{\infty} |h(k_1)| \cdot \sum_{k_2=0}^{\infty} |h(k_2)| \cdot \sum_{\ell_1=0}^{\infty} |h(\ell_1)| \cdot \sum_{\ell_2=0}^{\infty} |h(\ell_2)| \cdot \sum_{t=r}^{N} 4 \cdot C_e \le 4 \cdot C_e \cdot C_H^4 (N-r)$$

which proves (2B.5) of the lemma. The proof of (2B.6) is analogous and simpler. ■

Corollary to Lemma 2B.1. Let

$$w(t) = \sum_{k=0}^{\infty} \alpha_t(k)e(t-k), \qquad v(t) = \sum_{k=0}^{\infty} \beta_t(k)e(t-k)$$

Then

$$E\left\|\sum_{t=r}^{N} w(t)v(t) - Ew(t)v(t)\right\|^2 \le C \cdot C_w^2 \cdot C_v^2 \cdot (N-r)$$

$$C_w = \sum_{k=0}^{\infty} \sup_t |\alpha_t(k)|, \qquad C_v = \sum_{k=0}^{\infty} \sup_t |\beta_t(k)| \quad \blacksquare$$

Lemma 2B.2. Let

$$R_r^N = \sup_{\theta \in D_\theta} \left\| \sum_{t=r}^{N} s_\theta(t)s_\theta^T(t) - Es_\theta(t)s_\theta^T(t) \right\| \tag{2B.9}$$

Then

$$E(R_r^N)^2 \le C(N-r) \tag{2B.10}$$

Proof of Lemma 2B.2. First note the following fact: If

$$\varphi = \sum_{k=0}^{\infty} a(k)z(k) \tag{2B.11}$$

where $\{a(k)\}$ is a sequence of deterministic matrices, such that

$$\sum_{k=0}^{\infty} \|a(k)\| \le C_a$$

and $\{z(k)\}$ is a sequence of random vectors such that

$$E|z(k)|^2 \le C_z$$

then

$$\begin{aligned} E|\varphi|^2 &= \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \operatorname{tr}[a(k)Ez(k)z^T(\ell)a^T(\ell)] \\ &\le \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \|a(k)\| \cdot [E|z(k)|^2]^{1/2} \cdot [E|z(\ell)|^2]^{1/2} \cdot \|a(\ell)\| \\ &\le C_z \cdot \left[\sum_{k=0}^{\infty} \|a(k)\|\right]^2 \le C_z \cdot C_a^2 \end{aligned} \tag{2B.12}$$

Here the first inequality is Schwarz's inequality.

We now have

$$\begin{aligned} R_\theta(N,r) &= \sum_{t=r}^{N} [s_\theta(t)s_\theta^T(t) - Es_\theta(t)s_\theta^T(t)] \\ &= \sum_{t=r}^{N} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} g_\theta(k)[v(t-k)v^T(t-\ell) - Ev(t-k)v^T(t-\ell)]g_\theta^T(\ell) \\ &\quad + \sum_{t=r}^{N} \sum_{k=0}^{\infty} g_\theta(k)v(t-k)w_\theta^T(t) + \sum_{t=r}^{N} \sum_{\ell=0}^{\infty} w_\theta(t)v^T(t-\ell)g_\theta^T(\ell) \end{aligned} \tag{2B.13}$$

This gives

$$\sup_{\theta} \|R_\theta(N,r)\| \le \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \sup_{\theta} \|g_\theta(k)\| \cdot \|g_\theta(\ell)\| \cdot \|S_r^N(k,\ell)\| + 2C_w \cdot \sum_{k=0}^{\infty} \sup_{\theta} \|g_\theta(k)\| \, \|\tilde{S}_r^N(k)\| \tag{2B.14}$$

with S_r^N and $\tilde{S}_r^N$ defined by

$$S_r^N(k,\ell) = \sum_{t=r}^{N} [v(t-m)v^T(t-n) - Ev(t-m)v^T(t-n)]$$

$$\tilde{S}_r^N(k) = \sum_{t=r}^{N} [v(t-k)w_\theta^T(t) + w_\theta(t)v^T(t-k)]$$

Since $G_\theta(q)$ is a uniformly stable family of filters,

$$\sup_{\theta} \|g_\theta(k)\| \le \bar{g}(k), \qquad \sum_{k=1}^{\infty} \bar{g}(k) = C_G < \infty$$

Applying (2B.11) and (2B.12) together with Lemma 2B.1 to (2B.14) gives

$$E[\sup_{\theta}\|R_\theta(N,r)\|]^2 \le 2 \cdot C_G^4 \cdot 4 \cdot C_e \cdot C_H^4 \cdot (N-r) + 2 \cdot C_G^2 \cdot 4 \cdot C_w^2 \cdot C_H^2 \cdot (N-r) \le C \cdot (N-r)$$

which proves Lemma 2B.2. ■

We now turn to the proof of Theorem 2B.1. Denote

$$r(t,\theta) = s_\theta(t)s_\theta^T(t) - Es_\theta(t)s_\theta^T(t) \tag{2B.15}$$

and let

$$R_r^N = \sup_{\theta \in D_\theta} \|R_\theta(N,r)\| \tag{2B.16}$$

with $R_\theta(N,r)$ defined by (2B.13). According to Lemma 2B.2,

$$E\left(\frac{1}{N^2} R_1^{N^2}\right)^2 \le \left(\frac{1}{N^2}\right)^2 \cdot C \cdot N^2 \le \frac{C}{N^2}$$

Chebyshev's inequality (I.19) gives

$$P\left(\frac{1}{N^2} R_1^{N^2} > \varepsilon\right) \le \frac{1}{\varepsilon^2} E(R_1^{N^2})^2$$

Hence

$$\sum_{k=1}^{\infty} P\left(\frac{1}{k^2} R_1^{k^2} > \varepsilon\right) \le C \cdot \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty$$

which, via Borel–Cantelli's lemma [see (I.18)], implies that

$$\frac{1}{k^2} R_1^{k^2} \to 0, \qquad \text{w.p. 1,} \qquad \text{as } k \to \infty \tag{2B.17}$$

Now suppose that

$$\sup_{N^2 \le k \le (N+1)^2} \frac{1}{k} R_1^k$$

is obtained for $k = k_N$ and $\theta = \theta_N$. Hence

$$\begin{aligned}\sup_{N^2 \le k \le (N+1)^2} \frac{1}{k} R_1^k &= \frac{1}{k_N}\left|\sum_{t=1}^{k_N} r(t,\theta_N)\right| \\ &\le \frac{1}{k_N}\left|\sum_{t=1}^{N^2} r(t,\theta_N)\right| + \frac{1}{k_N}\left|\sum_{t=N^2+1}^{k_N} r(t,\theta_N)\right| \qquad (2B.18) \\ &\le \frac{1}{k_N} \cdot R_1^{N^2} + \frac{1}{k_N} \cdot R_{N^2+1}^{k_N}\end{aligned}$$

Since $k_N \ge N^2$, the first term on the right side of (2B.18) tends to zero w.p.1 in view of (2B.17). For the second one we have, using Lemma 2B.2,

$$E\left|\frac{1}{k_N} R_{N^2+1}^{k_N}\right|^2 \le \frac{1}{k_N^2} \cdot C \cdot [k_N - N^2 - 1] \le \frac{1}{N^4} \cdot C \cdot ((N+1)^2 - N^2 - 1) \le \frac{C}{N^3}$$

which, using Chebyshev's inequality (I.19) and the Borel–Cantelli lemma as before, shows that also the second term of (2B.18) tends to zero w.p. 1. Hence

$$\sup_{N^2 \le k \le (N+1)^2} \frac{1}{k} R_1^k \to 0, \qquad \text{w.p. 1,} \qquad \text{as } N \to \infty \qquad (2B.19)$$

which proves the theorem.

Corollary to Theorem 2B.1. Suppose that the conditions of the theorem hold, but that (2.86) is weakened to

$$E[e(t)|e(t-1), \ldots, e(0)] = 0, \qquad E[e^2(t)|e(t-1), \ldots, e(0)] = \lambda$$

$$E|e(t)|^4 \le C$$

Then the theorem still holds. [That is: $\{e(t)\}$ need not be white noise; it is sufficient that it is a martingale difference.]

Proof: Independence was used only in Lemma 2B.1. It is easy to see that this lemma holds also under the weaker conditions.

APPENDIX 2C: COVARIANCE FORMULAS

For several calculations we need expressions for variances and covariances of signals in ARMA descriptions. These are basically given by the inverse formulas

$$R_s(\tau) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_s(\omega) e^{i\omega\tau}\, d\omega \qquad (2C.1a)$$

$$R_{sv}(\tau) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_{sv}(\omega) e^{i\omega\tau}\, d\omega \qquad (2C.1b)$$

With the expressions according to Theorem 2.2 for the spectra, (2C.1a) takes the form

$$R_s(\tau) = \frac{\lambda}{2\pi}\int_{-\pi}^{\pi}\left|\frac{C(e^{i\omega})}{A(e^{i\omega})}\right|^2 e^{i\omega\tau}\,d\omega = [z = e^{i\omega}]$$

$$= \frac{\lambda}{2\pi}\oint \frac{C(z)C(1/z)}{A(z)A(1/z)}\,z^{\tau-1}\,dz \tag{2C.2}$$

for an ARMA process. The last integral is a complex integral around the unit circle, which could be evaluated using residue calculus. Åström, Jury, and Agniel, (1970) (see also Åström, 1970, Ch. 5) have derived an efficient algorithm for computing (2C.2) for $\tau = 0$. It has the following form:

$$A(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_n, \qquad C(z) = c_0 z^n + c_1 z^{n-1} + \cdots + c_n$$

Let $a_i^n = a_i$ and $c_i^n = c_i$ and define a_i^k, c_i^k recursively by

$$a_i^{n-k} = \frac{a_0^{n-k+1}a_i^{n-k+1} - a_{n-k+1}^{n-k+1}a_{n-k+1-i}^{n-k+1}}{a_0^{n-k+1}}$$

$$c_i^{n-k} = \frac{a_0^{n-k+1}c_i^{n-k+1} - c_{n-k+1}^{n-k+1}a_{n-k+1-i}^{n-k+1}}{a_0^{n-k+1}}$$

$$i = 0, 1, \ldots, n-k, \qquad k = 1, 2, \ldots, n$$

Then for (2C.2)

$$R_s(0) = \frac{1}{a_0}\sum_{k=0}^{n}\frac{(c_k^k)^2}{a_0^k} \tag{2C.3}$$

An explicit expression for the variance of a second-order ARMA process

$$y(t) + a_1 y(t-1) + a_2 y(t-2) = e(t) + c_1 e(t-1) + c_2 e(t-2)$$

$$Ee^2(t) = 1 \tag{2C.4}$$

is

Var $y(t) =$

$$\frac{(1+a_2)(1+(c_1)^2+(c_2)^2) - 2a_1c_1(1+c_2) - 2c_2(a_2 - (a_1)^2 + (a_2)^2)}{(1-a_2)(1-a_1+a_2)(1+a_1+a_2)} \tag{2C.5}$$

To find $R_y(\tau)$ and the cross covariances $R_{ye}(\tau)$ by hand calculations in simple examples, the easiest approach is to multiply (2C.4) by $e(t)$, $e(t-1)$, $e(t-2)$, $y(t)$, $y(t-1)$, and $y(t-2)$ and take expectation. This gives six equations for the six variables $R_{ye}(\tau)$, $R_y(\tau)$, $\tau = 0, 1, 2$. Note that $R_{ye}(\tau) = 0$ for $\tau < 0$.

3

SIMULATION, PREDICTION, AND CONTROL

The system descriptions given in Chapter 2 can be used for a variety of design problems related to the true system. In this chapter we shall discuss some such uses. The purpose of this is twofold. First, the idea of how to predict future output values will turn out to be most essential for the development of identification methods. The expressions provided in Section 3.2 will therefore be instrumental for the further discussion in this book. Second, by illustrating different uses of system descriptions, we will provide some insights into what is required for such descriptions to be adequate for their intended uses. A leading idea of our framework for identification will be that the effort spent in developing a model of a system must be related to the application it is going to be used for. Throughout the chapter we assume that the system description is given in the form (2.93):

$$y(t) = G(q)u(t) + H(q)e(t) \tag{3.1}$$

3.1 SIMULATION

The most basic use of a system description is to simulate the system's response to various input scenarios. This simply means that an input sequence $u^*(t)$, $t = 1, 2, \ldots, N$, chosen by the user is applied to (3.1) to compute the undisturbed output

$$y^*(t) = G(q)u^*(t), \qquad t = 1,2,\ldots,N \tag{3.2}$$

This is the output that the system would produce had there been no disturbances, according to the description (3.1). To evaluate the disturbance influence, a random-number generator (in the computer) is used to produce a sequence of numbers $e^*(t)$, $t = 1, 2, \ldots, N$, that can be considered as a realization of a white-noise stochastic process with variance λ. Then the disturbance is calculated as

$$v^*(t) = H(q)e^*(t) \tag{3.3}$$

By suitably presenting $y^*(t)$ and $v^*(t)$ to the user, an idea of the system's response to $\{u^*(t)\}$ can be formed.

This way of experimenting on the model (3.1) rather than on the actual, physical process to evaluate its behavior under various conditions has become widely used in engineering practice of all fields and no doubt reflects the most common use of mathematical descriptions. To be true, models used in, say, flight simulators or nuclear power station training simulators are of course far more complex than (3.1), but they still follow the same general idea (see also Chapter 5).

3.2 PREDICTION

We shall start by discussing how future values of $v(t)$ can be predicted in case it is described by

$$v(t) = H(q)e(t) = \sum_{k=0}^{\infty} h(k)e(t-k) \tag{3.4}$$

For (3.4) to be meaningful, we assume that H is stable; that is,

$$\sum_{k=0}^{\infty} |h(k)| < \infty \tag{3.5}$$

Invertibility of the Noise Model

A crucial property of (3.4), which we will impose, is that it should be *invertible*; that is, if $v(s)$, $s \le t$, are known, then we shall be able to compute $e(t)$ as

$$e(t) = \tilde{H}(q)v(t) = \sum_{k=0}^{\infty} \tilde{h}(k)v(t-k) \tag{3.6}$$

with

$$\sum_{k=0}^{\infty} |\tilde{h}(k)| < \infty$$

How can we determine the filter $\tilde{H}(q)$ from $H(q)$? The following lemma gives the answer.

Lemma 3.1. Consider $\{v(t)\}$ defined by (3.4) and assume that the filter H is stable. Let

$$H(z) = \sum_{k=0}^{\infty} h(k)z^{-k} \tag{3.7}$$

and assume that the function $1/H(z)$ is analytic in $|z| \geq 1$:

$$\frac{1}{H(z)} = \sum_{k=0}^{\infty} \tilde{h}(k) z^{-k} \tag{3.8}$$

Define the filter $H^{-1}(q)$ by

$$H^{-1}(q) = \sum_{k=0}^{\infty} \tilde{h}(k) q^{-k} \tag{3.9}$$

Then $\tilde{H}(q) = H^{-1}(q)$ satisfies (3.6).

Remark. That (3.8) exists for $|z| \geq 1$ also means that the filter $H^{-1}(q)$ is stable. For convenience, we shall then say that $H(q)$ is an *inversely stable* filter.

Proof. From (3.7) and (3.8) it follows that

$$1 = \sum_{k=0}^{\infty} \sum_{s=0}^{\infty} h(k)\tilde{h}(s) z^{-(k+s)} = [k + s = \ell] = \sum_{\ell=0}^{\infty} \sum_{k=0}^{\ell} h(k)\tilde{h}(\ell - k) z^{-\ell}$$

which implies that

$$\sum_{k=0}^{\ell} h(k)\tilde{h}(\ell - k) = \begin{cases} 1, & \text{if } \ell = 0 \\ 0, & \text{if } \ell \neq 0 \end{cases} \tag{3.10}$$

Now let $\{v(t)\}$ be defined by (3.4) and consider

$$\begin{aligned} \sum_{k=0}^{\infty} \tilde{h}(k) v(t-k) &= \sum_{k=0}^{\infty} \tilde{h}(k) \sum_{s=0}^{\infty} h(s) e(t-k-s) \\ &= \sum_{k=0}^{\infty} \sum_{s=0}^{\infty} \tilde{h}(k) h(s) e(t-k-s) = [k+s=\ell] \\ &= \sum_{\ell=0}^{\infty} \left[\sum_{k=0}^{\infty} \tilde{h}(k) h(\ell - k) \right] e(t-\ell) = e(t) \end{aligned}$$

according to (3.10), which proves the lemma. ■

Note: The lemma shows that the properties of *the filter* $H(q)$ are quite analogous to those of *the function* $H(z)$. It is not a triviality that the inverse filter $H^{-1}(q)$ can be drived by inverting the function $H(z)$; hence the formulation of the result as a lemma. However, all similar relationships between $H(q)$ and $H(z)$ will also hold, and from a practical point of view it will be useful to switch freely between the filter and its z-transform. See also Problem 3D.1.

The lemma shows that the inverse filter (3.6) in a natural way relates to the original filter (3.4). In view of its definition, we shall also write

$$H^{-1}(q) = \frac{1}{H(q)} \tag{3.11}$$

for this filter. All that is needed is that the function $1/H(z)$ be analytic in $|z| \geq 1$; that is, it has no poles on or outside the unit circle. We could also phrase the condition as $H(z)$ must have no zeros on or outside the unit circle. This ties in very

nicely with the spectral factorization result (see Section 2.3) according to which, for rational strictly positive spectra, we can always find a representation $H(q)$ with these properties.

Example 3.1 A moving average process

Suppose that

$$v(t) = e(t) + ce(t-1) \tag{3.12}$$

That is,

$$H(q) = 1 + cq^{-1}$$

According to (2.85), this process is a *moving average of order 1*, MA(1). Then

$$H(z) = 1 + cz^{-1} = \frac{z+c}{z}$$

has a pole in $z = 0$ and a zero in $z = -c$, which is inside the unit circle if $|c| < 1$. If so, the inverse filter is determined as

$$H^{-1}(z) = \frac{1}{H(z)} = \frac{1}{1+cz^{-1}} = \sum_{k=0}^{\infty} (-c)^k z^{-k}$$

and $e(t)$ is recovered from (3.12) as

$$e(t) = \sum_{k=0}^{\infty} (-c)^k v(t-k)$$

■

One-step-ahead Prediction of v

Suppose now that we have observed $v(s)$ for $s \le t-1$ and that we want to predict the value of $v(t)$ based on these observations. We have, since H is monic,

$$v(t) = \sum_{k=0}^{\infty} h(k)e(t-k) = e(t) + \sum_{k=1}^{\infty} h(k)e(t-k) \tag{3.13}$$

Now, the knowledge of $v(s)$, $s \le t-1$ implies the knowledge of $e(s)$, $s \le t-1$, in view of (3.6). The second term of (3.13) is therefore known at time $t-1$. Let us denote it, provisionally, by $m(t-1)$:

$$m(t-1) = \sum_{k=1}^{\infty} h(k)e(t-k)$$

Suppose that $\{e(t)\}$ are identically distributed, and let the probability distribution of $e(t)$ be denoted by $f_e(x)$:

$$P(x \le e(t) \le x + \Delta x) \approx f_e(x)\Delta x$$

This distribution is independent of the other values of $e(s)$, $s \ne t$, since $\{e(t)\}$ is a sequence of independent random variables. What we can say about $v(t)$ at time

$t-1$ is consequently that the probability that $v(t)$ assumes a value between $m(t-1)+x$ and $m(t-1)+x+\Delta x$ is $f_e(x)\,\Delta x$. This could also be phrased as

> the (posterior) probability density function of $v(t)$, given observations up to time $t-1$, is $f_v(x) = f_e(x - m(t-1))$.

Formally, these calculations can be written as

$$f_v(x)\Delta x = P(x \le v(t) \le x + \Delta x | v_{-\infty}^{t-1}) = P(x \le m(t-1) + e(t) \le x + \Delta x)$$
$$= P(x - m(t-1) \le e(t) \le x + \Delta x - m(t-1)) = f_e(x - m(t-1))\Delta x$$

Here $P(A|v_{-\infty}^{t-1})$ means the conditional probability of the event A, given $v_{-\infty}^{t-1}$.

This is the most complete statement that can be made about $v(t)$ at time $t-1$. Often we just give one value that characterizes this probability distribution and hence serves as a *prediction* of $v(t)$. This could be chosen as the value for which the PDF $f_e(x - m(t-1))$ has its maximum, the most probable value of $v(t)$, which also is called the *maximum a posteriori (MAP)* prediction. We shall, however, mostly work with the mean value of the distribution in question, the *conditional expectation* of $v(t)$ denoted by $\hat{v}(t|t-1)$. Since the variable $e(t)$ has zero mean, we have

$$\hat{v}(t|t-1) = m(t-1) = \sum_{k=1}^{\infty} h(k)e(t-k) \tag{3.14}$$

It is easy to establish that the conditional expectation also minimizes the mean-square error of the prediction error:

$$\min_{\tilde{v}(t)} \quad E(v(t) - \tilde{v}(t))^2 \Rightarrow \tilde{v}(t) = \hat{v}(t|t-1)$$

where the minimization is carried out over all functions $\tilde{v}(t)$ of $v_{-\infty}^{t-1}$. See Problem 3D.3.

Let us find a more convenient expression for (3.14). We have, using (3.6) and (3.11),

$$\hat{v}(t|t-1) = \left[\sum_{k=1}^{\infty} h(k)q^{-k}\right]e(t) = [H(q) - 1]e(t)$$
$$= \frac{H(q)-1}{H(q)}\,v(t) = [1 - H^{-1}(q)]v(t) = \sum_{k=1}^{\infty} -\tilde{h}(k)v(t-k) \tag{3.15}$$

Applying $H(q)$ to both sides gives the alternative expression

$$H(q)\hat{v}(t|t-1) = [H(q) - 1]v(t) = \sum_{k=1}^{\infty} h(k)v(t-k) \tag{3.16}$$

Example 3.2 A moving average process

Consider the process (3.12). Then (3.16) shows that the predictor is calculated as

$$\hat{v}(t|t-1) + c\hat{v}(t-1|t-2) = cv(t-1) \tag{3.17}$$

Alternatively we can determine $H^{-1}(q)$ from Example 3.1 and use (3.15):

$$\hat{v}(t|t-1) = -\sum_{k=1}^{\infty} (-c)^k v(t-k)$$

■

Example 3.3 An autoregressive process

Consider a process

$$v(t) = \sum_{k=0}^{\infty} a^k e(t-k), \qquad |a| < 1$$

Then

$$H(z) = \sum_{k=0}^{\infty} a^k z^{-k} = \frac{1}{1-az^{-1}}$$

which gives

$$H^{-1}(z) = 1 - az^{-1}$$

and the predictor, according to (3.15),

$$\hat{v}(t|t-1) = av(t-1) \tag{3.18}$$

■

One-step-ahead Prediction of y

Consider the description (3.1), and assume that $y(s)$ and $u(s)$ are known for $s \le t-1$. Since

$$v(s) = y(s) - G(q)u(s) \tag{3.19}$$

this means that also $v(s)$ are known for $s \le t-1$. We would like to predict the value

$$y(t) = G(q)u(t) + v(t)$$

based on this information. Clearly, the conditional expectation of $y(t)$, given the information in question, is

$$\begin{aligned}\hat{y}(t|t-1) &= G(q)u(t) + \hat{v}(t|t-1)\\ &= G(q)u(t) + [1 - H^{-1}(q)]v(t)\\ &= G(q)u(t) + [1 - H^{-1}(q)][y(t) - G(q)u(t)]\end{aligned}$$

using (3.15) and (3.19), respectively. Collecting the terms gives

$$\hat{y}(t|t-1) = H^{-1}(q)G(q)u(t) + [1 - H^{-1}(q)]y(t) \tag{3.20}$$

or

$$H(q)\hat{y}(t|t-1) = G(q)u(t) + [H(q) - 1]y(t) \tag{3.21}$$

Remember that these expressions are shorthand notation for expansions. For example, let $\{\ell(k)\}$ be defined by

$$\frac{G(z)}{H(z)} = \sum_{k=1}^{\infty} \ell(k) z^{-k} \tag{3.22}$$

[This expansion exists for $|z| \geq 1$ if $H(z)$ has no zeros and $G(z)$ no poles in $|z| \geq 1$.] Then (3.20) means that

$$\hat{y}(t|t-1) = \sum_{k=1}^{\infty} \ell(k) u(t-k) + \sum_{k=1}^{\infty} -\tilde{h}(k) y(t-k) \tag{3.23}$$

Unknown Initial Conditions

In the reasoning so far we have made use of the assumption that the whole data record from time minus infinity to $t-1$ is available. Indeed, in the expression (3.20) as in (3.23) all these data appear explicitly. In practice, however, it is usually the case that only data over the interval $[0, t-1]$ are known. The simplest thing would then be to replace the unknown data by zero (say) in (3.23):

$$\hat{y}(t|t-1) \approx \sum_{k=1}^{t} \ell(k) u(t-k) + \sum_{k=1}^{t} -\tilde{h}(k) y(t-k) \tag{3.24}$$

One should realize that this is now only an approximation of the actual conditional expectation of $y(t)$, given data over $[0, t-1]$. The exact prediction involves time-varying filter coefficients and can be computed using the Kalman filter [see (4.91)]. For most practical purposes, (3.24) will, however, give a satisfactory solution. The reason is that the coefficients $\{\ell(k)\}$ and $\{\tilde{h}(k)\}$ typically decay exponentially with k (see Problem 3G.1).

The Prediction Error

From (3.20) and (3.1), we find that the prediction error $y(t) - \hat{y}(t|t-1)$ is given by

$$y(t) - \hat{y}(t|t-1) = -H^{-1}(q)G(q)u(t) + H^{-1}(q)y(t) = e(t) \tag{3.25}$$

The variable $e(t)$ thus represents that part of the output $y(t)$ that cannot be predicted from past data. For this reason it is also called the *innovation* at time t.

k-step-ahead Prediction of y (*)

Having treated the problem of one-step-ahead prediction in some detail, it is easy to generalize to the following problem: Suppose that we have observed $v(s)$ for $s \leq t$ and that we want to predict the value $v(t+k)$. We have

$$v(t+k) = \sum_{\ell=0}^{\infty} h(\ell)e(t+k-\ell) = \sum_{\ell=0}^{k-1} h(\ell)e(t+k-\ell) + \sum_{\ell=k}^{\infty} h(\ell)e(t+k-\ell) \tag{3.26}$$

Let us define

$$\bar{H}_k(q) = \sum_{\ell=0}^{k-1} h(\ell)q^{-\ell}, \qquad \tilde{H}_k(q) = \sum_{\ell=k}^{\infty} h(\ell)q^{-\ell+k} \tag{3.27}$$

The second sum of (3.26) is known at time t, while the first sum is independent of what has happened up to time t and has zero mean. The conditional mean of $v(t+k)$, given $v_{-\infty}^{t}$ is thus given by

$$\hat{v}(t+k|t) = \sum_{\ell=k}^{\infty} h(\ell)e(t+k-\ell) = \tilde{H}_k(q)e(t) = \tilde{H}_k(q)\cdot H^{-1}(q)v(t)$$

This expression is the k-step-ahead predictor of v.

Now suppose that we have measured $y_{-\infty}^{t}$ and know $u_{-\infty}^{t+k-1}$ and would like to predict $y(t+k)$. We have, as before,

$$y(t+k) = G(q)u(t+k) + v(t+k)$$

which gives

$$\begin{aligned}\hat{y}(t+k|y_{-\infty}^{t}, u_{-\infty}^{t+k-1}) &\triangleq \hat{y}(t+k|t) = G(q)u(t+k) + \hat{v}(t+k|t)\\ &= G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)v(t)\\ &= G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]\end{aligned} \tag{3.28}$$

Introduce

$$\begin{aligned}W_k(q) &\triangleq 1 - q^{-k}\tilde{H}_k(q)H^{-1}(q) = [H(q) - q^{-k}\tilde{H}_k(q)]H^{-1}(q)\\ &= \bar{H}_k(q)H^{-1}(q)\end{aligned} \tag{3.29}$$

Then simple manipulation on (3.28) gives

$$\hat{y}(t+k|t) = W_k(q)G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)y(t) \tag{3.30}$$

or, using the first equality in (3.29),

$$\hat{y}(t|t-k) = W_k(q)G(q)u(t) + [1 - W_k(q)]y(t) \tag{3.31}$$

This expression, together with (3.27) and (3.29), defines the k-step-ahead predictor for y. Notice that this predictor can also be viewed as a one-step-ahead predictor associated with the model

$$y(t) = G(q)u(t) + W_k^{-1}(q)e(t) \tag{3.32}$$

The prediction error is obtained from (3.30) as

$$\begin{aligned} e_k(t+k) &\triangleq y(t+k) - \hat{y}(t+k|t) = -W_k(q)G(q)u(t+k) \\ &\quad + [q^k - \tilde{H}_k(q)H^{-1}(q)]y(t) \\ &= W_k(q)[y(t+k) - G(q)u(t+k)] = W_k(q)H(q)e(t+k) \\ &= \bar{H}_k(q)e(t+k) \end{aligned} \tag{3.33}$$

Here we used (3.29) in the second and fourth equalities. According to (3.27), $\bar{H}_k(q)$ is a polynomial in q^{-1} of order $k-1$. Hence the prediction error is a moving average of $e(t+k), \ldots, e(t+1)$.

The Multivariable Case (*)

For a multivariable system description (3.1) (or 2.88), we define the $p \times p$ matrix filter $H^{-1}(q)$ as

$$H^{-1}(q) = \sum_{k=0}^{\infty} \tilde{h}(k)q^{-k}$$

Here $\tilde{h}(k)$ are the $p \times p$ matrices defined by the expansion of the matrix function

$$[H(z)]^{-1} = \sum_{k=0}^{\infty} \tilde{h}(k)z^{-k} \tag{3.34}$$

This expansion can be interpreted entrywise in the matrix $[H(z)]^{-1}$ (formed by standard manipulations for matrix inversion). It exists for $|z| \geq 1$ provided the function $\det H(z)$ has no zeros in $|z| \geq 1$. With $H^{-1}(q)$ thus defined, all calculations and formulas given previously are valid also for the multivariable case.

3.3 OBSERVERS

In many cases in systems and control theory, one does not work with a full description of the properties of disturbances as in (3.1). Instead a noise-free or "*deterministic*" model is used:

$$y(t) = G(q)u(t) \tag{3.35}$$

In this case one probably keeps in the back of one's mind, though, that (3.35) is not really the full story about the input–output properties.

The description (3.35) can of course also be used for "computing," "guessing," or "predicting" future values of the output. The lack of noise model, however, leaves several possibilities for how this can best be done. The concept of *observers* is a key issue for these calculations. This concept is normally discussed in terms of state-space representations of (3.35) (see Section 4.3); see, for example Luenberger (1971) or Åström and Wittenmark (1984). But it can equally well be introduced for the input–output form (3.35).

An Example

Let

$$G(z) = b \sum_{k=1}^{\infty} (a)^{k-1} z^{-k} = \frac{bz^{-1}}{1 - az^{-1}} \tag{3.36}$$

This means that the input–output relationship can be represented either as

$$y(t) = b \sum_{k=1}^{\infty} (a)^{k-1} u(t-k) \tag{3.37}$$

that is

$$y(t) = \frac{bq^{-1}}{1 - aq^{-1}} u(t)$$

or as

$$(1 - aq^{-1}) y(t) = bq^{-1} u(t)$$

i.e.

$$y(t) - ay(t-1) = bu(t-1) \tag{3.38}$$

Now, if we are given the description (3.35) and (3.36) together with data $y(s)$, $u(s)$, $s \leq t-1$, and are asked to produce a "guess" or to "calculate" what $y(t)$ might be, we could use either

$$\hat{y}(t|t-1) = b \sum_{k=1}^{\infty} (a)^{k-1} u(t-k) \tag{3.39}$$

or

$$\hat{y}(t|t-1) = ay(t-1) + bu(t-1) \tag{3.40}$$

As long as the data and the system description are correct, there would also be no difference between (3.39) and (3.40); they are both "observers" (in our setting "predictors" would be a more appropriate term) for the system. The choice between them would be carried out by the designer in terms of how vulnerable they are to imperfections in data and descriptions. For example, if input–output data are lacking prior to time $s = 0$, then (3.39) suffers from an error that decays like a^t (effect of wrong initial conditions), whereas (3.40) is still correct for $t \geq 1$. On the other hand, (3.39) is unaffected by measurement errors in the output, whereas such errors are directly transferred to the prediction in (3.40). From the discussion of Section 3.2, it should be clear that, if (3.35) is complemented with a noise model as in (3.1), then the choice of predictor becomes unique (cf. Problem 3E.3). This follows since the conditional mean of the output, computed according to the assumed noise model, is a uniquely defined quantity.

A Family of Predictors for (3.35)

The example (3.36) showed that the choice of predictor could be seen as a trade-off between sensitivity with respect to output measurement errors and rapidly decaying effects of erroneous initial conditions. To introduce design variables for this trade-off, choose a filter $W(q)$ such that

$$W(q) = 1 + \sum_{\ell=k}^{\infty} w_\ell q^{-\ell} \tag{3.41}$$

Apply it to both sides of (3.35):

$$W(q)y(t) = W(q)G(q)u(t)$$

which means that

$$y(t) = [1 - W(q)]y(t) + W(q)G(q)u(t)$$

In view of (3.41), the right side of this expression depends only on $y(s)$, $s \le t - k$, and $u(s)$, $s \le t - 1$. Based on that information, we could thus produce a "guess" or prediction of $y(t)$ as

$$\hat{y}(t|t-k) = [1 - W(q)]y(t) + W(q)G(q)u(t) \tag{3.42}$$

The trade-off considerations for the choice of W would then be:

1. Select $W(q)$ so that both W and WG have rapidly decaying filter coefficients in order to minimize the influence of erroneous initial conditions.
2. Select $W(q)$ so that measurement imperfections in $y(t)$ are maximally attenuated.

(3.43)

The later issue can be illuminated in the frequency domain: Suppose that $y(t) = y_M(t) + v(t)$, where $y_M(t) = G(q)u(t)$ is the useful signal and $v(t)$ is a measurement error. Then the prediction error according to (3.42) is

$$\varepsilon(t) = y(t) - \hat{y}(t|t-k) = W(q)v(t) \tag{3.44}$$

The spectrum of this error is, according to Theorem 2.2,

$$\Phi_\varepsilon(\omega) = |W(e^{i\omega})|^2 \Phi_v(\omega) \tag{3.45}$$

where $\Phi_v(\omega)$ is the spectrum of v. The problem is thus to select W, subject to (3.41), such that the error spectrum (3.45) has an acceptable size and suitable shape.

A comparison with the k-step prediction case of Section 3.2 shows that the expression (3.42) is identical to (3.31) with $W(q) = W_k(q)$. It is clear that the qualification of a complete noise model in (3.1) allows us to analytically compute the filter W in accordance with aspect 2. This was indeed what we did in Section 3.2. However, aspect 1 was neglected there, since we assumed all past data to be available. Normally, as we pointed out, this aspect is also less important.

Figure 3.1 The predictor filter.

Fundamental Role of the Predictor Filter

It turns out that for most uses of system descriptions it is the predictor form (3.20), or as in (3.31) and (3.42), that is more important than the description (3.1) or (3.35) itself. We use (3.31) and (3.42) to predict, or "guess," future outputs; we use it for control design (as we shall see) to regulate the predicted output, and so on. Now, (3.31) and (3.42) are just linear filters into which sequences $\{u(t)\}$ and $\{y(t)\}$ are fed, and that produce $\hat{y}(t|t-k)$ as output. The thoughts that the designer had when he or she selected this filter are immaterial once it is put to use: *The filter is the same whether $W = W_k$ was chosen as a trade-off (3.43) or computed from H as in (3.27) and (3.29).* The noise model H in (3.1) is from this point of view just an alibi for determining the predictor. This is the viewpoint we are going to adopt. The predictor filter is the fundamental system description (Figure 3.1). Our rationale for arriving at the filter is secondary. This also means that the difference between a "stochastic system" (3.1) and a "deterministic" one (3.35) is not fundamental. Nevertheless, we find it convenient to use the description (3.1) as the basic system description. It is in a one-to-one correspondence with the one-step-ahead predictor (3.20) (see Problem 3D.2) and relates more immediately to traditional system descriptions.

3.4 CONTROL (*)

The problem of control is to make the output of a system behave in a desired fashion by properly selecting the input sequence. It is quite natural that a description of the system in one form or another will be required in order to adequately design such a control mechanism. In this section we shall briefly describe how the description (3.1) can be used for some typical control design problems.

Classical Lead–Lag Compensation

"Classical" control theory as developed by Bode, Nichols, and Nyquist, among others, in the 1940s is based on some graphical representation of the frequency function $G(e^{i\omega})$. A typical approach is to use a feedback law

$$u(t) = F(q)[r(t) - y(t)] \tag{3.46}$$

where $r(t)$ is the desired value of the output: the *set point* or the *reference signal.* The closed-loop system then becomes (neglecting the disturbances)

$$y(t) = \frac{F(q)G(q)}{1 + F(q)G(q)} r(t) \tag{3.47}$$

and the selection of F is made so that the compensated frequency function

$$F(e^{i\omega})G(e^{i\omega})$$

has the desired properties, which typically are checked in their graphical representation. For stability of (3.47), the values of the frequency function around certain frequencies, like the crossover frequency, where

$$|F(e^{i\omega})G(e^{i\omega})| = 1$$

are especially important. It follows that a good knowledge of $G(e^{i\omega})$ at these frequencies will be of great importance for successful application of this technique. See any elementary textbook on control design for a further discussion (e.g., Kuo, 1982).

Minimum Variance Control

The idea behind minimum variance control of (3.1) is to select the input so that the output has the smallest possible variance around zero mean value. Suppose that there is a time delay of k time units from the input to the output; that is,

$$G(q) = \sum_{\ell=k}^{\infty} g(\ell)q^{-\ell}, \qquad g(k) \neq 0 \tag{3.48}$$

This means that the choice of $u(t)$ will affect $y(t+k)$ but no output prior to this. Now we can write [see (3.33)]

$$y(t+k) = \hat{y}(t+k|t) + e_k(t+k)$$

where $e_k(t+k)$ depends only on $e(t+k), \ldots, e(t+1)$. This term cannot therefore be affected by $u(t)$. Also, $e_k(t+k)$ and $\hat{y}(t+k|t)$ are independent, because $\{e(t+k)\ldots e(t+1)\}$ are independent of $u(s)$, $s \leq t$, and $y(s)$, $s \leq t$ [see also (3.30) to (3.33)]. Hence

$$E\,y(t+k)^2 = E\,\hat{y}(t+k|t)^2 + E\,e_k(t+k)^2$$

Clearly, then, the variance of $y(t+k)$ is minimized if we can make $\hat{y}(t+k|t) = 0$. From (3.29) and (3.31), we see that the regulator

$$u(t) = -\frac{\bar{H}_k(q)H^{-1}(q)}{W_k(q)q^kG(q)}\,y(t) \tag{3.49}$$

will achieve this. Equation (3.49) consequently defines the *minimum variance regulator*. Note that the calculation of (3.49) is essentially the same as the calculation of the k-step-ahead predictor.

In the special case of $k = 1$, the preceding expression simplifies to

$$u(t) = -\frac{H(q) - 1}{G(q)}\,y(t) \tag{3.50}$$

We may also remark that these regulators are meaningful only if $q^kG(q)$ is inversely stable [i.e., if $G(z)$ has no zeros on or outside the unit circle].

Shaping of the Noise Spectrum

Consider now, for simplicity, the case of one delay in the system [i.e., $k = 1$ in (3.48)]. Suppose that we would like the output of the system (3.1) to behave like

$$y(t) = R(q)e(t) \tag{3.51}$$

where R is a monic filter.

Clearly, with $R(q) = 1$ we would obtain minimum variance control. Other R's will give a higher output variance, but could still be very well motivated, allowing for more reasonable trade-offs between control effort and output properties.

It is easy to verify that the regulator

$$u(t) = \frac{R(q) - H(q)}{G(q)R(q)}y(t) \tag{3.52}$$

achieves (3.51) when inserted into (3.1). This is thus the sought regulator [in case $G(z)$ has no zeros on or outside the unit circle].

Synthesis of a Given Closed-loop System

Consider again the case $k = 1$, and assume that the control design objective is that

$$y(t) = R(q)r(t) + e(t) \tag{3.53}$$

where $r(t)$ is the reference signal and $R(q)$ is some prespecified desired closed-loop transfer function. Since $y(t) = \hat{y}(t|t-1) + e(t)$, we achieve (3.53) if the predicted value obeys

$$\hat{y}(t|t-1) = R(q)r(t) \tag{3.54}$$

Try the regulator

$$u(t) = F_1(q)r(t) - F_2(q)y(t) \tag{3.55a}$$

and insert it into (3.20). This gives

$$\hat{y}(t|t-1) = H^{-1}(q)G(q)F_1(q)r(t) + [1 - H^{-1}(q) - H^{-1}(q)G(q)F_2(q)]y(t)$$

We see that the choices

$$F_1(q) = \frac{R(q)H(q)}{G(q)} \tag{3.55b}$$

and

$$F_2(q) = \frac{H(q) - 1}{G(q)} \tag{3.55c}$$

will give the desired behavior (3.54). Note that, for $F_1(q)$ to be causal, $R(q)$ must have at least one delay. Also, this regulator is realistic only if $G(z)$ has no zeros outside the unit circle.

3.5 SUMMARY

Starting from the representation

$$y(t) = G(q)u(t) + H(q)e(t)$$

we have derived an expression for the one-step-ahead prediction of $y(t)$ [i.e., the best "guess" of $y(t)$ given $u(s)$ and $y(s)$, $s \le t - 1$]. This expression is given by

$$\hat{y}(t|t-1) = H^{-1}(q)G(q)u(t) + [1 - H^{-1}(q)]y(t) \tag{3.56}$$

We also derived a corresponding k-step-ahead predictor (3.31). We pointed out that one can arrive at such predictors also through deterministic observer considerations, not relying on a noise model H. We have stressed that the bottom line in most uses of a system description is how these predictions actually are computed; the underlying noise assumptions are merely vehicles for arriving at the predictors. The discussion of Chapters 2 and 3 can thus be viewed as a methodology for "guessing" future system outputs.

We have also illustrated some uses of the system descriptions and the predictor expressions for control design.

It should be noted that calculations such as (3.56) involved in determining the predictors and regulators are typically performed with greater computational efficiency once they are applied to transfer functions G and H with more specific structures. This will be illustrated in the next chapter.

3.6 BIBLIOGRAPHY

Prediction and control are standard textbook topics. Accounts of the k-step-ahead predictor and associated control problems can be found in Åström (1970) and Åström and Wittenmark (1984). Prediction is treated in detail in, for example, Anderson and Moore (1979) and Box and Jenkins (1970). An early account of this theory is Whittle (1963).

Prediction theory was developed by Kolmogorov (1941a), Wiener (1949), Kalman (1960), and Kalman and Bucy (1961). The hard part in these problems is indeed to find a suitable representation of the disturbance. Once we arrive at (3.1) via spectral factorization, or at its time-varying counterpart via the Riccati equation [see (4.92) and Problem 4G.3], the calculation of a reasonable predictor is, as demonstrated here, easy. Note, however (as pointed out in Problem 2E.3), that for non-Gaussian processes normally only the second-order properties can be adequately described by (3.1), which consequently is too simple a representation to accommodate more complex noise structures. The calculations carried out in Section 3.2 are given in Åström (1970) for the case where G and H are rational with the same denominators. Rissanen and Barbosa (1969) have given expressions for the prediction in input–output models of this kind when the lack of knowledge of the infinite past is treated properly [i.e., when the ad hoc solution (3.24) is not accepted]. The result is, of course, a time-varying predictor.

Calculations of regulators as in (3.48) to (3.55) may look quite similar in the multivariable case. Structural problems are, however, then quite nontrivial. See Kucera (1979) and Pernebo (1981) for treatments of such problems.

3.7 PROBLEMS

3G.1. Suppose that the transfer function $G(z)$ is rational and that its poles are all inside $|z| < \mu$, where $\mu < 1$. Show that

$$|g(k)| \leq c \cdot \mu^k$$

where $g(k)$ is defined as in (2.16).

3G.2. Let $A(q)$ and $B(q)$ be two monic stable and inversely stable filters. Show that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |A(e^{i\omega})|^2 \cdot |B(e^{i\omega})|^2 \, d\omega \geq 1$$

with equality only if $A(q) = 1/B(q)$.

3E.1. Let

$$H(q) = 1 - 1.1q^{-1} + 0.3q^{-2}$$

Compute $H^{-1}(q)$ as an explicit infinite expansion.

3E.2. Determine the 3-step-ahead predictors for

$$y(t) = \frac{1}{1 - aq^{-1}} e(t)$$

and

$$y(t) = (1 + cq^{-1})e(t)$$

respectively. What are the variances of the associated prediction errors?

3E.3. Show that if (3.35) and (3.36) are complemented with the noise model $H(q) = 1$ then (3.39) is the natural predictor, whereas the noise model

$$H(q) = \sum_{k=0}^{\infty} (a^k)q^{-k}$$

leads to the predictor (3.40).

3E.4. Let $e(t)$ have the distribution

$$e(t) = \begin{cases} 1, & \text{w.p. } \frac{1}{2} \\ -0.5, & \text{w.p. } \frac{1}{4} \\ -1.5, & \text{w.p. } \frac{1}{4} \end{cases}$$

Let

$$v(t) = H(q)e(t)$$

and let $\hat{v}(t|t-1)$ be defined as in the text. What is the most probable value (MAP) of $v(t)$ given the information $\hat{v}(t|t-1)$? What is the probability that $v(t)$ will assume a value between $\hat{v}(t|t-1) - \frac{1}{4}$ and $\hat{v}(t|t-1) + \frac{1}{4}$?

3T.1. Suppose that $A(q)$ is inversely stable and monic. Show that $A^{-1}(q)$ is monic.

3T.2. Suppose the measurement error spectrum of v in (3.44) and (3.45) is given by

$$\Phi_v(\omega) = \lambda |R(e^{i\omega})|^2$$

for some monic stable and inversely stable filter $R(q)$. Find the filter W, subject to (3.41) with $k = 1$, that minimizes

$$\bar{E}\varepsilon^2(t)$$

Hint: Use Problem 3G.2.

3T.3. Consider the system description of Problem 2E.4:

$$x(t+1) = fx(t) + w(t)$$
$$y(t) = hx(t) + v(t)$$

(x scalar). Assume that $\{v(t)\}$ is white Gaussian noise with variance R_2 and that $\{w(t)\}$ is a sequence of independent variables with

$$w(t) = \begin{cases} 1, & \text{w.p. } 0.05 \\ -1, & \text{w.p. } 0.05 \\ 0, & \text{w.p. } 0.9 \end{cases}$$

Determine a monic filter $W(q)$ such that the predictor

$$\hat{y}(t) = (1 - W(q))y(t)$$

minimizes

$$\bar{E}(y(t) - \hat{y}(t))^2$$

What can be said about

$$E(y(t)|y_{-\infty}^{t-1})?$$

3T.4. Consider the noise description

$$v(t) = e(t) + ce(t-1), \qquad |c| > 1, \qquad Ee^2(t) = \lambda \tag{3.57}$$

Show that $e(t)$ cannot be reconstructed from v^t by a causal, stable filter. However, show that $e(t)$ can be computed from v_{t+1}^{∞} by an anticausal, stable filter. Thus construct a stable, anticausal predictor for $v(t)$ given $v(s)$, $s \geq t+1$.

Determine a noise $\bar{v}(t)$ with the same second-order properties as $v(t)$, such that

$$\bar{v}(t) = \bar{e}(t) + c^*\bar{e}(t-1), \qquad |c^*| < 1, \qquad E\bar{e}^2(t) = \lambda^* \tag{3.58}$$

Show that $\bar{v}(t)$ can be predicted from $\bar{v}^{t-1}$ by a stable, causal predictor. [Measuring just second-order properties of the noise, we cannot distinguish between (3.57) and and (3.58). However, when $e(t)$ in (3.57) is a physically well defined quantity (although not measured by us), we may be interested in which one of (3.57) and (3.58) has generated the noise. See Benveniste, Goursat, and Ruget (1980).]

3D.1. In the chapter we have freely multiplied, added, subtracted, and divided by transfer-function operators $G(q)$ and $H(q)$. Division was formalized and justified by Lemma 3.1 and (3.11). Justify similarly addition and multiplication.

3D.2. Suppose a one-step-ahead predictor is given as

$$\hat{y}(t|t-1) = L_1(q)u(t-1) + L_2(q)y(t-1)$$

Calculate the system description (3.1) from which this predictor was derived.

3D.3. Consider a stochastic process $\{v(t)\}$ and let

$$\hat{v}(t) = E(v(t)|v^{t-1})$$

Define

$$e(t) = v(t) - \hat{v}(t)$$

Let $\tilde{v}(t)$ be an arbitrary function of v^{t-1}. Show that

$$E(v(t) - \tilde{v}(t))^2 \geq E\, e^2(t)$$

Hint: Use $Ex^2 = E_z E(x^2|z)$.

4

MODELS OF LINEAR TIME-INVARIANT SYSTEMS

A model of a system is a description of (some of) its properties, suitable for a certain purpose. The model need not be a true and accurate description of the system, nor need the user have to believe so, in order to serve its purpose.

System identification is the subject of constructing or selecting models of dynamical systems to serve certain purposes. As we noted in Chapter 1, a first step is to determine a class of models within which the search for the most suitable model is to be conducted. In this chapter we shall discuss such classes of models for linear time-invariant systems.

4.1 LINEAR MODELS AND SETS OF LINEAR MODELS

A linear time-invariant model is specified, as we saw in Chapter 2, by the impulse response $\{g(k)\}_1^\infty$, the spectrum $\Phi_v(\omega) = \lambda|H(e^{i\omega})|^2$ of the additive disturbance, and, possibly, the probability density function (PDF) of the disturbance $e(t)$. A complete model is thus given by

$$y(t) = G(q)u(t) + H(q)e(t) \tag{4.1}$$

$$f_e(\cdot), \text{ the PDF of } e$$

with

$$G(q) = \sum_{k=1}^{\infty} g(k)q^{-k}, \qquad H(q) = 1 + \sum_{k=1}^{\infty} h(k)q^{-k} \tag{4.2}$$

A particular model thus corresponds to specification of the three functions G, H, and f_e. It is in most cases impractical to make this specification by enumerating the infinite sequences $\{g(k)\}$, $\{h(k)\}$ together with the function $f_e(x)$. Instead one chooses to work with structures that permit the specification of G and H in terms of a finite number of numerical values. Rational transfer functions and finite-dimensional state-space descriptions are typical examples of this. Also, most often the PDF f_e is not specified as a function, but described in terms of a few numerical characteristics, typically the first and second moments:

$$\begin{aligned} E\,e(t) &= \int x f_e(x)\,dx = 0 \\ E\,e^2(t) &= \int x^2 f_e(x)\,dx = \lambda \end{aligned} \tag{4.3}$$

It is also common to assume that $e(t)$ is Gaussian, in which case the PDF is entirely specified by (4.3). The specification of (4.1) in terms of a finite number of numerical values, or coefficients, has another and most important consequence for the purposes of system identification. Quite often it is not possible to determine these coefficients a priori from knowledge of the physical mechanisms that govern the system's behavior. Instead the determination of all or some of them must be left to estimation procedures. This means that the coefficients in question enter the model (4.1) as *parameters to be determined.* We shall generally denote such parameters by the vector θ, and thus have a model description

$$y(t) = G(q,\theta)u(t) + H(q,\theta)e(t) \tag{4.4a}$$

$$f_e(x,\theta),\ \text{the PDF of } e(t);\ \{e(t)\}\ \text{white noise} \tag{4.4b}$$

The parameter vector θ the ranges over a subset of $\mathbf{R}^d$, where d is the dimension of θ:

$$\theta \in D_{\mathcal{M}} \subset \mathbf{R}^d \tag{4.5}$$

Notice that (4.4) to (4.5) no longer is *a* model; it is a *set of models,* and it is for the estimation procedure to select that member in the set that appears to be most suitable for the purpose in question. [One may sometimes loosely talk about "the model (4.4)," but this is abuse of notation from a formal point of view.] Using (3.20), we can compute the one-step-ahead prediction for (4.4). Let it be denoted by $\hat{y}(t|\theta)$ to emphasize its dependence on θ. We thus have

$$\hat{y}(t|\theta) = H^{-1}(q,\theta)G(q,\theta)u(t) + [1 - H^{-1}(q,\theta)]y(t) \tag{4.6}$$

This predictor form does not depend on $f_e(x,\theta)$. In fact, as we stressed in Section 3.3, we could very well arrive at (4.6) by considerations that are not probabilistic.

Then the specification (4.4b) does not apply. We shall use the term *predictor models* for models that just specify G and H as in (4.4a) or in the form (4.6). Similarly, *probabilistic models* will signify descriptions (4.4) that give a complete characterization of the probabilistic properties of the system. A parametrized set of models like (4.6) will be called a *model structure* and will be denoted by $\mathcal{M}$. The particular model associated with the parameter value θ will be denoted by $\mathcal{M}(\theta)$. (A formal definition is given in Section 4.5.)

In the following three sections, different ways of describing (4.4a) in terms of θ (i.e., different ways of parametrizing the model set) will be discussed. A formalization of the concepts of model sets, parametrizations, model structures, and uniqueness of parametrization will then be given in Section 4.5, while questions of identifiability are discussed in Section 4.6.

4.2 A FAMILY OF TRANSFER-FUNCTION MODELS

Perhaps the most immediate way of parametrizing G and H is to represent them as rational functions and let the parameters be the numerator and denominator coefficients. In this section we shall describe various ways of carrying out such parametrizations. Such model structures are also known as black-box models.

Equation Error Model Structure

Probably the most simple input–output relationship is obtained by describing it as a linear difference equation:

$$y(t) + a_1 y(t-1) + \cdots + a_{n_a} y(t-n_a) = b_1 u(t-1) + \cdots + b_{n_b} u(t-n_b) + e(t) \tag{4.7}$$

Since the white-noise term $e(t)$ here enters as a direct error in the difference equation, the model (4.7) is often called an *equation error model* (structure). The adjustable parameters are in this case

$$\theta = [a_1 \quad a_2 \ldots a_{n_a} \quad b_1 \ldots b_{n_b}]^T \tag{4.8}$$

If we introduce

$$A(q) = 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a}$$

and

$$B(q) = b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b}$$

we see that (4.7) corresponds to (4.4) with

$$G(q,\theta) = \frac{B(q)}{A(q)}, \qquad H(q,\theta) = \frac{1}{A(q)} \tag{4.9}$$

Remark. It may seem annoying to use q as an argument of $A(q)$, being a polynomial in q^{-1}. The reason for this is, however, simply to be consistent with the conventional definition of the z-transform; see (2.17). ■

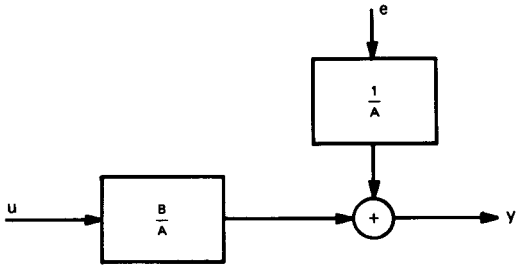

Figure 4.1 The ARX model structure.

We shall also call the model (4.7) an ARX model, where AR refers to the autoregressive part $A(q)y(t)$ and X to the extra input $B(q)u(t)$ (called the exogeneous variable in econometrics). In the special case where $n_a = 0$, $y(t)$ is modeled as a finite impulse response (FIR). Such model sets are particularly common in signal-processing applications.

The signal flow can be depicted as in Figure 4.1. From that picture we see that the model (4.7) is perhaps not the most natural one from a physical point of view: the white noise is assumed to go through the denominator dynamics of the system before being added to the output. Nevertheless, the equation error model set has a very important property that makes it a prime choice in many applications: The predictor defines a linear regression.

Linear Regressions

Let us compute the predictor for (4.7). Inserting (4.9) into (4.6) gives

$$\hat{y}(t|\theta) = B(q)u(t) + [1 - A(q)]y(t) \tag{4.10}$$

Clearly, this expression could have more easily been derived directly from (4.7). Let us reiterate the view expressed in Section 3.3: Without a stochastic framework, the predictor (4.10) is a natural choice if the term $e(t)$ in (4.7) is considered to be "insignificant" or "difficult to guess." It is thus perfectly natural to work with the expression (4.10) also for "deterministic" models.

Now introduce the vector

$$\varphi(t) = [-y(t-1) \dots -y(t-n_a)\ u(t-1) \dots u(t-n_b)]^T \tag{4.11}$$

Then (4.10) can be rewritten as

$$\hat{y}(t|\theta) = \theta^T\varphi(t) = \varphi^T(t)\theta \tag{4.12}$$

This is the important property of (4.7) that we alluded to previously. The predictor is a scalar product between a known data vector $\varphi(t)$ and the parameter vector θ. Such a model is called a *linear regression* in statistics, and the vector $\varphi(t)$ is known as the *regression vector*. It is of importance since powerful and simple estimation methods can be applied for the determination of θ.

In case some coefficients of the polynomials A and B are known, we arrive at linear regressions of the form

$$\hat{y}(t|\theta) = \varphi^T(t)\theta + \mu(t) \tag{4.13}$$

where $\mu(t)$ is a known term. See Problem 4E.1 and also (5.34). The estimation of θ in linear regressions will be treated in Section 7.3. See also Appendix II.

ARMAX Model Structure

The basic disadvantage with the simple model (4.7) is the lack of adequate freedom in describing the properties of the disturbance term. We could add flexibility to that by describing the equation error as a moving average of white noise. This gives the model

$$\begin{aligned} &y(t) + a_1 y(t-1) + \cdots + a_{n_a} y(t-n_a) \\ &= b_1 u(t-1) + \cdots + b_{n_b} u(t-n_b) + e(t) + c_1 e(t-1) + \cdots + c_{n_c} e(t-n_c) \end{aligned} \tag{4.14}$$

With

$$C(q) = 1 + c_1 q^{-1} + \cdots + c_{n_c} q^{-n_c}$$

it can be rewritten

$$A(q)y(t) = B(q)u(t) + C(q)e(t) \tag{4.15}$$

and clearly corresponds to (4.4) with

$$G(q,\theta) = \frac{B(q)}{A(q)}, \qquad H(q,\theta) = \frac{C(q)}{A(q)} \tag{4.16}$$

where now

$$\theta = [a_1 \ldots a_{n_a} b_1 \ldots b_{n_b} c_1 \ldots c_{n_c}]^T \tag{4.17}$$

In view of the moving average (MA) part $C(q)e(t)$, the model (4.15) will be called ARMAX. The ARMAX model has become a standard tool in control and econometrics for both system description and control design. A version with an enforced integration in the system description is the ARIMA(X) model (I for integration, with or without the X-variable u), which is useful to describe systems with slow disturbances; see Box and Jenkins (1970). It is obtained by replacing $y(t)$ in (4.15) by $\Delta y(t) = y(t) - y(t-1)$ and is further discussed in Section 14.6.

Pseudolinear Regressions

The predictor for (4.15) is obtained by inserting (4.16) into (4.6). This gives

$$\hat{y}(t|\theta) = \frac{B(q)}{C(q)}u(t) + \left[1 - \frac{A(q)}{C(q)}\right]y(t)$$

or

$$C(q)\hat{y}(t|\theta) = B(q)u(t) + [C(q) - A(q)]y(t) \tag{4.18}$$

This means that the prediction is obtained by filtering u and y through a filter with denominator dynamics determined by $C(q)$. To start it up at time $t = 0$ requires knowledge of

$$\begin{aligned} &\hat{y}(0|\theta) \dots \hat{y}(-n_c + 1|\theta) \\ &y(0) \dots y(-n^* + 1), \qquad n^* = \max(n_c, n_a) \\ &u(0) \dots u(-n_b + 1) \end{aligned}$$

If these are not available, they can be taken as zero, in which case the prediction differs from the true one with an error that decays as $c \cdot \mu^t$, where μ is the maximum magnitude of the zeros of $C(z)$. It is also possible to start the recursion at time $\max(n^*, n_b)$ and include the unknown initial conditions $\hat{y}(k|\theta)$, $k = 1, \dots, n_c$, in the vector θ.

The predictor (4.18) can be rewritten in formal analogy with (4.12) as follows. Adding $[1 - C(q)]\hat{y}(t|\theta)$ to both sides of (4.18) gives

$$\hat{y}(t|\theta) = B(q)u(t) + [1 - A(q)]y(t) + [C(q) - 1][y(t) - \hat{y}(t|\theta)] \tag{4.19}$$

Introduce the prediction error

$$\varepsilon(t,\theta) = y(t) - \hat{y}(t|\theta)$$

and the vector

$$\begin{aligned} \varphi(t,\theta) = [&-y(t-1) \dots -y(t-n_a)\, u(t-1) \dots u(t-n_b) \\ &\varepsilon(t-1,\theta) \dots \varepsilon(t-n_c,\theta)]^T \end{aligned} \tag{4.20}$$

Then (4.19) can be rewritten as

$$\hat{y}(t|\theta) = \varphi^T(t,\theta)\theta \tag{4.21}$$

Notice the similarity with the linear regression (4.12). The equation (4.21) itself is, however, no linear regression, due to the nonlinear effect of θ in the vector $\varphi(t, \theta)$. To stress the kinship to (4.12), we shall call it a *pseudolinear regression.*

Other Equation-Error-Type Model Structures

Instead of modeling the equation error in (4.7) as a moving average, as we did in (4.14), it can of course be described as an autoregression. This gives a model set

$$A(q)y(t) = B(q)u(t) + \frac{1}{D(q)}e(t) \tag{4.22}$$

with

$$D(q) = 1 + d_1 q^{-1} + \cdots + d_{n_d} q^{-n_d}$$

which, analogously to the previous terminology, could be called ARARX. More generally, we could use an ARMA description of the equation error, leading to an "ARARMAX" structure

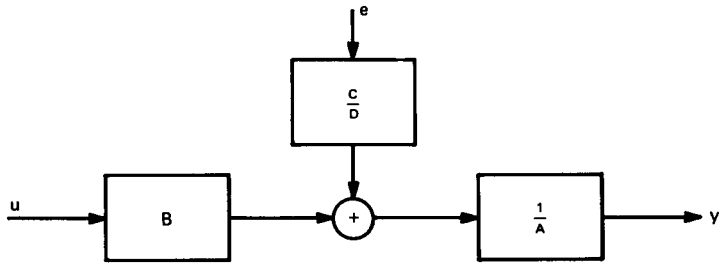

Figure 4.2 The equation error model family: The model structure (4.23).

$$A(q)y(t) = B(q)u(t) + \frac{C(q)}{D(q)}e(t) \tag{4.23}$$

which of course contains (4.7), (4.15), and (4.22) as special cases. This would thus form the family of equation-error-related model sets, and is depicted in Figure 4.2. The relationship to (4.4) as well as expressions for the predictions are straightforward.

Output Error Model Structure

The equation error model structures all correspond to descriptions where the transfer functions G and H have the polynomial A as a common factor in the denominators. See Figure 4.2. From a physical point of view it may seem more natural to parametrize these transfer functions independently.

If we suppose that the relation between input and undisturbed output w can be written as a linear difference equation, and that the disturbances consist of white measurement noise, then we obtain the following description:

$$w(t) + f_1 w(t-1) + \cdots + f_{n_f} w(t-n_f) = b_1 u(t-1) + \cdots + b_{n_b} u(t-n_b) \tag{4.24a}$$

$$y(t) = w(t) + e(t) \tag{4.24b}$$

With

$$F(q) = 1 + f_1 q^{-1} + \cdots + f_{n_f} q^{-n_f}$$

we can write the model as

$$y(t) = \frac{B(q)}{F(q)}u(t) + e(t) \tag{4.25}$$

The signal flow of this model is shown in Figure 4.3.

We call (4.25) an *output error (OE) model* (structure). The parameter vector to be determined is

$$\theta = [b_1\ b_2 \ldots b_{n_b}\ f_1 f_2 \ldots f_{n_f}]^T \tag{4.26}$$

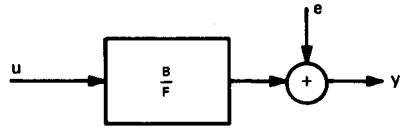

Figure 4.3 The output error model structure.

Since $w(t)$ in (4.24a) is never observed, it should rightly carry an index θ, since it is constructed from u using (4.24a). That is,

$$w(t,\theta) + f_1 w(t-1,\theta) + \cdots + f_{n_f} w(t-n_f,\theta) = b_1 u(t-1) + \cdots + b_{n_b} u(t-n_b) \tag{4.27}$$

Comparing with (4.4), we find that $H(q,\theta) = 1$, which gives the natural predictor

$$\hat{y}(t|\theta) = \frac{B(q)}{F(q)} u(t) = w(t,\theta) \tag{4.28}$$

With the aid of the vector

$$\varphi(t,\theta) = [u(t-1) \ldots u(t-n_b) \quad -w(t-1,\theta) \ldots -w(t-n_f,\theta)]^T \tag{4.29}$$

this can be rewritten as

$$\hat{y}(t|\theta) = \varphi^T(t,\theta)\theta \tag{4.30}$$

which is in formal agreement with the ARMAX-model predictor (4.21). Note that in (4.29) the $w(t-1, \theta)$ are not observed, but, using (4.28), they can be computed: $w(t-k, \theta) = \hat{y}(t-k|\theta)$, $k = 1, 2, \ldots, n_f$.

Box–Jenkins Model Structure

A natural development of the output error model (4.25) is to further model the properties of the output error. Describing this as an ARMA model gives

$$y(t) = \frac{B(q)}{F(q)} u(t) + \frac{C(q)}{D(q)} e(t) \tag{4.31}$$

In a sense, this is the most natural finite-dimensional parametrization, starting from the descripton (4.4): the transfer functions G and H are independently parametrized as rational functions. The model set (4.31) was suggested and treated in Box and Jenkins (1970). This model also gives us the family of output-error-related models. See Figure 4.4 and compare with Figure 4.2. According to (4.6), the predictor for (4.31) is

$$\hat{y}(t|\theta) = \frac{D(q)B(q)}{C(q)F(q)} u(t) + \frac{C(q) - D(q)}{C(q)} y(t) \tag{4.32}$$

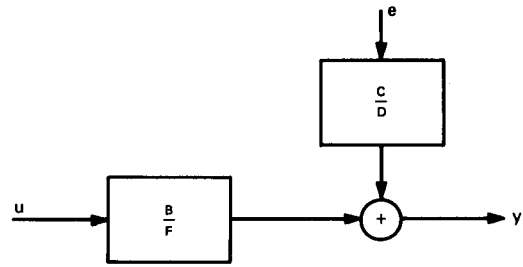

Figure 4.4 The BJ-model structure (4.31).

A General Family of Model Structures

The structures we have discussed in this section actually may give rise to 32 different model sets, depending on which of the five polynomials A, B, C, B, and F are used. (We have, however, only explicitly displayed six of these possibilities here.) Several of these model sets belong to the most commonly used ones in practice, and we have therefore reason to return to them both for explicit algorithms and for analytic results. For convenience, we shall therefore use a generalized model structure

$$A(q)y(t) = \frac{B(q)}{F(q)}u(t) + \frac{C(q)}{D(q)}e(t) \tag{4.33}$$

Sometimes the dynamics from u to y contains a delay of n_k samples, so some leading coefficients of B are zero; that is,

$$B(q) = b_{n_k}q^{-n_k} + b_{n_k+1}q^{-n_k-1} + \cdots + b_{n_k+n_b-1}q^{-n_k-n_b+1} = q^{-n_k}\bar{B}(q), \qquad b_{n_k} \neq 0$$

It may then be a good idea to explicitly display this delay by

$$A(q)y(t) = q^{-n_k}\frac{\bar{B}(q)}{F(q)}u(t) + \frac{C(q)}{D(q)}e(t) \tag{4.34}$$

For easier notation we shall, however, here mostly use $n_k = 1$ and (4.33). From expressions for (4.33) we can always derive the corresponding ones for (4.34) by replacing $u(t)$ by $u(t - n_k + 1)$.

The structure (4.33) is too general for most practical purposes. One or several of the five polynomials would be fixed to unity in applications. However, by developing algorithms and results for (4.33), we also cover all the special cases corresponding to more realistic model sets.

From (4.6) we know that the predictor for (4.33) is

$$\hat{y}(t|\theta) = \frac{D(q)B(q)}{C(q)F(q)}u(t) + \left[1 - \frac{D(q)A(q)}{C(q)}\right]y(t) \tag{4.35}$$

The common special cases of (4.33) are summarized in Table 4.1.

TABLE 4.1 Some Common Black-box SISO Models as Special Cases of (4.33)

Polynomials Used in (4.33)	Name of Model Structure
B	FIR (finite impulse response)
AB	ARX
ABC	ARMAX
AC	ARMA
ABD	ARARX
$ABCD$	ARARMAX
BF	OE (output error)
$BFCD$	BJ (Box–Jenkins)

A Pseudolinear Form for (4.35) (*)

The expression (4.35) can also be written as a recursion:

$$C(q)F(q)\hat{y}(t|\theta) = F(q)[C(q) - D(q)A(q)]y(t) + D(q)B(q)u(t) \tag{4.36}$$

From (4.36) we find that the prediction error

$$\varepsilon(t,\theta) = y(t) - \hat{y}(t|\theta)$$

can be written

$$\varepsilon(t,\theta) = \frac{D(q)}{C(q)}\left[A(q)y(t) - \frac{B(q)}{F(q)}u(t)\right] \tag{4.37}$$

It is convenient to introduce the auxiliary variables

$$w(t,\theta) = \frac{B(q)}{F(q)}u(t) \tag{4.38a}$$

and

$$v(t,\theta) = A(q)y(t) - w(t,\theta) \tag{4.38b}$$

Then

$$\varepsilon(t,\theta) = y(t) - \hat{y}(t|\theta) = \frac{D(q)}{C(q)}v(t,\theta) \tag{4.39}$$

Let us also introduce the "state vector"

$$\begin{aligned}\varphi(t,\theta) = [&-y(t-1),\ldots,-y(t-n_a),u(t-1),\ldots,u(t-n_b),\\ &-w(t-1,\theta),\ldots,-w(t-n_f,\theta),\varepsilon(t-1,\theta),\ldots,\varepsilon(t-n_c,\theta),\\ &-v(t-1,\theta),\ldots,-v(t-n_d,\theta)]^T\end{aligned} \tag{4.40}$$

With the parameter vector

$$\theta = [a_1 \ldots a_{n_a}\, b_1 \ldots b_{n_b} f_1 \ldots f_{n_f} c_1 \ldots c_{n_c}\, d_1 \ldots d_{n_d}]^T \tag{4.41}$$

and (4.40) we can give a convenient expression for the prediction. To find this, we proceed as follows: From (4.38) and (4.39) we obtain

$$w(t,\theta) = b_1 u(t-1) + \cdots + b_{n_b} u(t-n_b) - f_1 w(t-1,\theta) - \cdots - f_{n_f} w(t-n_f,\theta) \tag{4.42}$$

and

$$\varepsilon(t,\theta) = v(t,\theta) + d_1 v(t-1,\theta) + \cdots + d_{n_d} v(t-n_d,\theta) - c_1 \varepsilon(t-1,\theta) - \cdots - c_{n_c} \varepsilon(t-n_c,\theta) \tag{4.43}$$

Now inserting

$$v(t,\theta) = y(t) + a_1 y(t-1) + \cdots + a_{n_a} y(t-n_a) - w(t,\theta)$$

into (4.43) and substituting $w(t,\theta)$ with the expression (4.42), we find that

$$\varepsilon(t,\theta) = y(t) - \theta^T \varphi(t,\theta) \tag{4.44}$$

Hence

$$\hat{y}(t|\theta) = \theta^T \varphi(t,\theta) = \varphi^T(t,\theta)\theta \tag{4.45}$$

The two expressions, (4.36) and (4.45) can both be used for the calculation of the prediction. It should be noticed that the expressions simplify considerably in the special cases of the general model (4.33) that have been discussed in this section.

Continuous-time Black-box Models (*)

The linear system description could also be parametrized in terms of the continuous-time transfer function (2.22):

$$y(t) = G_c(p,\theta)u(t) \tag{4.46}$$

Adjustments to observed, sampled data could then be achieved either by solving the underlying differential equations or by applying an exact or approximate sampling procedure (2.24). In addition to obvious counterparts of the structures already discussed, two specific model sets should be mentioned. The first-order system model with a time delay

$$G_c(s,\theta) = \frac{Ke^{-s\tau_c}}{(s\tau + 1)}, \qquad \theta = (K, \tau_c, \tau) \tag{4.47}$$

has been much used in process industry applications. Orthonormal function series expansions

$$G_c(s,\theta) = \sum_{k=0}^{d-1} a_k f_k(s), \qquad \theta = (a_0, \cdots, a_{d-1}) \tag{4.48}$$

have been discussed in the early literature, and also more recently by Belanger (1985). For orthonormal functions, Laguerre polynomials appear to be a good choice:

$$f_k(s) = \sqrt{2\alpha}\,\frac{(s-\alpha)^k}{(s+\alpha)^{k+1}}$$

α being a time-scaling factor. Clearly, the model (4.46) can then be complemented with a model for the disturbance effects at the sampling instants as in (2.23).

Multivariable Case: Matrix Fraction Descriptions (*)

Let us now consider the case where the input $u(t)$ is an m-dimensional vector and the output $y(t)$ is a p-dimensional vector. Most of the ideas that we have described in this section have straightforward multivariable counterparts. The simplest case is the generalization of the equation error model set (4.7). We obtain

$$\begin{aligned} y(t) + A_1 y(t-1) + \cdots + A_{n_a} y(t-n_a) \\ = B_1 u(t-1) + \cdots + B_{n_b} u(t-n_b) + e(t) \end{aligned} \tag{4.49}$$

where the A_i are $p \times p$ matrices and the B_i are $p \times m$ matrices.

Analogous to (4.9), we may introduce the polynomials

$$\begin{aligned} A(q) &= I + A_1 q^{-1} + \cdots + A_{n_a} q^{-n_a} \\ B(q) &= B_1 q^{-1} + \cdots + B_{n_b} q^{-n_b} \end{aligned} \tag{4.50}$$

These are now *matrix polynomials* in q^{-1}, meaning that $A(q)$ is a matrix whose entries are polynomials in q^{-1}. We note that the system is still given by

$$y(t) = G(q,\theta)u(t) + H(q,\theta)e(t) \tag{4.51}$$

with

$$G(q,\theta) = A^{-1}(q)B(q), \qquad H(q,\theta) = A^{-1}(q) \tag{4.52}$$

The inverse $A^{-1}(q)$ of the matrix polynomial is interpreted and calculated in a straightforward way as discussed in connection with (3.34). Clearly, $G(q,\theta)$ will be a $p \times m$ matrix whose entries are rational functions of q^{-1} (or q). The factorization in terms of two matrix polynomials is also called a (left) *matrix fraction description* (MFD). A thorough treatment of such descriptions is given in Chapter 6 of Kailath (1980).

We have not yet discussed the *parametrization* of (4.49) (i.e., which elements of the matrices should be included in the parameter vector θ). This is a fairly subtle issue, which will be further discussed in Appendix 4A. An immediate analog of (4.8) could, however, be noted: Suppose all matrix entries in (4.49) (a total of $n_a \cdot p^2 + n_b \cdot p \cdot m$) are included in θ. We may then define the $[n_a \cdot p + n_b \cdot m] \times p$ matrix

$$\boldsymbol{\theta} = [A_1 A_2 \cdots A_{n_a} B_1 \cdots B_{n_b}]^T \tag{4.53}$$

and the $[n_a \cdot p + n_b \cdot m]$-dimensional column vector

$$\varphi(t) = \begin{bmatrix} -y(t-1) \\ \vdots \\ -y(t-n_a) \\ u(t-1) \\ \vdots \\ u(t-n_b) \end{bmatrix} \tag{4.54}$$

to rewrite (4.49) as

$$y(t) = \boldsymbol{\theta}^T \varphi(t) + e(t) \tag{4.55}$$

in obvious analogy with the linear regression (4.12). This can be seen as p different linear regressions, written on top of each other, all with the same regression vector.

When additional structure is imposed on the parametrization, it is normally no longer possible to use (4.55), since the different output components will not employ identical regression vectors. Then a d-dimensional column vector θ and a $p \times d$ matrix $\varphi^T(t)$ has to be formed so as to represent (4.49) as

$$y(t) = \varphi^T(t)\theta + e(t) \tag{4.56}$$

See Problems 4G.6 and 4E.12 for some more aspects on (4.55) and (4.56).

In light of the different possibilities for SISO systems, it is easy to visualize a number of variants for the MIMO case, like the vector difference equation (VDE)

$$y(t) + A_1 y(t-1) + \cdots + A_{n_a} y(t-n_a) = B_1 u(t-1) + \cdots + B_{n_b} u(t-n_b) + e(t) + C_1 e(t-1) + \cdots + C_{n_c} e(t-n_c) \tag{4.57a}$$

or

$$G(q,\theta) = A^{-1}(q)B(q), \qquad H(q,\theta) = A^{-1}(q)C(q) \tag{4.57b}$$

which is the natural extension of the ARMAX model. A multivariable Box–Jenkins model takes the form

$$G(q,\theta) = F^{-1}(q)B(q), \qquad H(q,\theta) = D^{-1}(q)C(q) \tag{4.58}$$

and so on. The parametrizations of these MFD-descriptions are discussed in Appendix 4A.

4.3 STATE-SPACE MODELS

In the state-space form the relationship between the input, noise, and output signals is written as a system of first-order differential or difference equations using an auxiliary state vector $x(t)$. This description of linear dynamical systems became an increasingly dominating approach after Kalman's (1960) work on prediction and linear quadratic control. For our purposes it is especially useful in that insights into physical mechanisms of the system can usually more easily be incorporated into state-space models than into the models described in Section 4.2.

Continuous-time Models Based on Physical Insight

For most physical systems it is easier to construct models with physical insight in continuous time than in discrete time, simply because most laws of physics (Newton's law of motion, relationships in electrical circuits, etc.) are expressed in continuous time. This means that modeling normally leads to a representation

$$\dot{x}(t) = F(\theta)x(t) + G(\theta)u(t) \tag{4.59}$$

Here F and G are matrices of appropriate dimensions ($n \times n$ and $n \times m$, respectively, for an n-dimensional state and an m-dimensional input). The overdot denotes differentiation with respect to (w.r.t) time t. Moreover, θ is a vector of parameters that typically correspond to unknown values of physical coefficients, material constants, and the like. The modeling is usually carried out in terms of state variables x that have physical significance (positions, velocities, etc.), and then the measured outputs will be known combinations of the states. Let $\eta(t)$ be the measurements that would be obtained with ideal, noise-free sensors:

$$\eta(t) = Hx(t) \tag{4.60}$$

Using p for the differentiation operator, (4.59) can be written

$$[pI - F(\theta)]x(t) = G(\theta)u(t)$$

which means that the transfer operator from u to η in (4.60) is

$$\begin{aligned} \eta(t) &= G_c(p,\theta)u(t) \\ G_c(p,\theta) &= H[pI - F(\theta)]^{-1}G(\theta) \end{aligned} \tag{4.61}$$

We have thus obtained a continuous-time transfer-function model of the system, as in (2.22), that is parametrized in terms of physical coefficients.

In reality, of course, some noise-corrupted version of $\eta(t)$ is obtained, resulting from both measurement imperfections and disturbances acting on (4.59). There are several different possibilities to describe these noise and disturbance effects. Here we first take the simplest approach. Other cases are discussed in (4.81) and (4.93) to (4.96), in Problem 4G.7, and in Section 14.5. Let the measurements be sampled at time instants $t = kT$, $k = 1, 2, \cdots$, and the disturbance effects at those time instants be $v_T(kT)$. Hence the measured output is

$$y(kT) = Hx(kT) + v_T(kT) = G_c(p,\theta)u(t) + v_T(kT) \tag{4.62}$$

Sampling the Transfer Function

As we discussed in Section 2.1, there are several ways of transporting $G_c(p,\theta)$ to a representation that is explicitly discrete time. Suppose that the input is constant over the sampling interval T as in (2.3):

$$u(t) = u_k = u(kT), \qquad kT \le t < (k+1)T \tag{4.63}$$

Then the differential equation (4.59) can easily be solved from $t = kT$ to $t = kT + T$, yielding

$$x(kT + T) = A_T(\theta)x(kT) + B_T(\theta)u(kT) \tag{4.64}$$

where

$$A_T(\theta) = e^{F(\theta)T} \tag{4.65a}$$

$$B_T(\theta) = \int_{\tau=0}^{T} e^{F(\theta)\tau} G(\theta)\, d\tau \tag{4.65b}$$

(See, e.g., Åström and Wittenmark, 1984.)

Introducing q for the forward shift of T time units, we can rewrite (4.64) as

$$[qI - A_T(\theta)]x(kT) = B_T(\theta)u(kT) \tag{4.66}$$

or

$$\eta(kT) = G_T(q, \theta)u(kT) \tag{4.67}$$

$$G_T(q, \theta) = H[qI - A_T(\theta)]^{-1}B_T(\theta) \tag{4.68}$$

Hence (4.62) can equivalently be given in the sampled-data form

$$y(t) = G_T(q, \theta)u(t) + v_T(t), \qquad t = T, 2T, 3T, \ldots \tag{4.69}$$

When (4.63) holds, no approximation is involved in this representation. Note, however, that in view of (4.65) $G_T(q, \theta)$ could be quite a complicated function of θ.

Example 4.1 DC servomotor

In this example we shall study a physical process, where we have some insight into the dynamic properties. Consider the dc motor depicted in Figure 4.5 with a block diagram in Figure 4.6. The input to this system is assumed to be the applied voltage, u, and the output the angle of the motor shaft, η. The relationship between applied voltage u and the resulting current i in the rotor circuit is given by the well-known relationship

$$u(t) = R_a i(t) + L_a \frac{di(t)}{dt} + s(t) \tag{4.70}$$

where $s(t)$ is the back electromotive force, due to the rotation of the armature circuit in the magnetic field:

$$s(t) = k_v \frac{d}{dt}\eta(t)$$

The current i gives a turning torque of

$$T_a(t) = K_a \cdot i(t)$$

on the motor shaft, which is also affected by a torque $T_\ell(t)$ from the load. Newton's law then gives

$$J \frac{d^2}{dt^2}\eta(t) = T_a(t) - T_\ell(t) - f \frac{d}{dt}\eta(t) \tag{4.71}$$

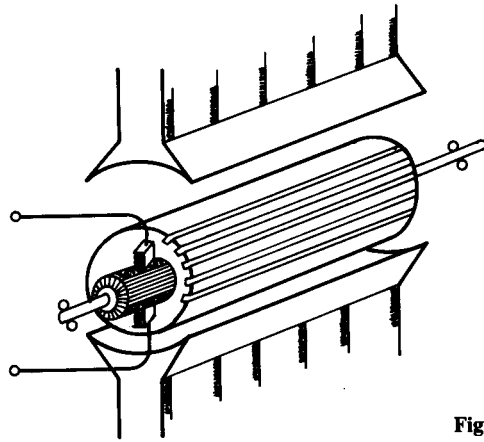

Figure 4.5 The dc motor.

where J is the moment of inertia of the rotor plus load and f represents viscous friction. Assuming that the inductance of the armature circuit can be neglected, $L_a \approx 0$, the preceding equations can be summarized in state-space form as

$$\frac{d}{dt}x(t) = \begin{bmatrix} 0 & 1 \\ 0 & -1/\tau \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ \frac{\beta}{\tau} \end{bmatrix} u(t) + \begin{bmatrix} 0 \\ \frac{\gamma'}{\tau} \end{bmatrix} T_\ell(t) \tag{4.72}$$

$$\eta(t) = [1 \quad 0]\, x(t)$$

with

$$x(t) = \begin{bmatrix} \eta(t) \\ \frac{d}{dt}\eta(t) \end{bmatrix}$$

$$\tau = \frac{JR_a}{fR_a + k_a k_v}, \qquad \beta = \frac{k_a}{fR_a + k_a k_v}, \qquad \gamma' = -\frac{R_a}{fR_a + k_a k_v}$$

Assume now that the torque T_ℓ is identically zero. To determine the dynamics of the motor, we now apply a piecewise constant input and sample the output with the sampling interval T. The state equation (4.72) can then be described by

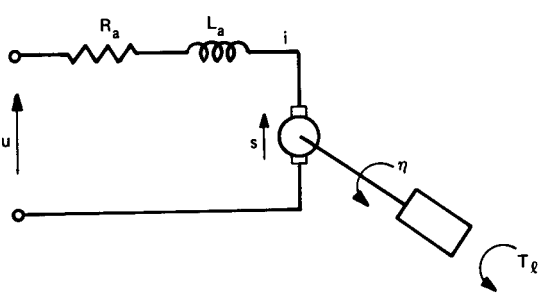

Figure 4.6 Block diagram of the dc motor.

$$x(t+T) = A_T(\theta)x(t) + B_T(\theta)u(t) \tag{4.73}$$

where

$$\theta = \begin{bmatrix} \tau \\ \beta \end{bmatrix}$$

and, according to (4.65),

$$A_T(\theta) = \begin{bmatrix} 1 & \tau(1-e^{-T/\tau}) \\ 0 & e^{-T/\tau} \end{bmatrix}$$
$$B_T(\theta) = \begin{bmatrix} \beta(\tau e^{-T/\tau} - \tau + T) \\ \beta(1-e^{-T/\tau}) \end{bmatrix} \tag{4.74}$$

Also assume that $y(t)$, the actual measurement of the angle $\eta(t)$, is made with a certain error $v(t)$:

$$y(t) = \eta(t) + v(t) \tag{4.75}$$

This error is mainly caused by limited accuracy (e.g., due to the winding of a potentiometer) and can be described as a sequence of independent random variables with zero mean and known variance R_2 (computed from the truncation error in the measurement), provided the measurements are not too frequent. We thus have a model

$$y(t) = G_T(q,\theta)u(t) + v(t)$$

with $v(t)$ being white noise. The natural predictor is thus

$$\hat{y}(t|\theta) = G_T(q,\theta)u(t) = [1 \quad 0][qI - A_T(\theta)]^{-1}B_T(\theta) \tag{4.76}$$

This predictor is parametrized using only *two* parameters β and τ. Notice that if we used our physical insight to conclude only that the system is of second order we would use, say, a second-order ARX or OE model containing *four* adjustable parameters. As we shall see, using fewer parameters has some positive effects on the estimation procedure: the variance of the parameter estimates will decrease. The price is, however, not insignificant. The predictor (4.76) is a far more complicated function of its two parameters than the corresponding ARX or OE model of its four parameters. ■

Equations (4.64) and (4.62) constitute a standard discrete-time state-space model. For simplicity we henceforth take $T = 1$ and drop the corresponding index. We also introduce an arbitrary parametrization of the matrix that relates x to η: $H = C(\theta)$. We thus have

$$x(t+1) = A(\theta)x(t) + B(\theta)u(t) \tag{4.77a}$$
$$y(t) = C(\theta)x(t) + v(t) \tag{4.77b}$$

corresponding to

$$y(t) = G(q,\theta)u(t) + v(t) \tag{4.78}$$
$$G(q,\theta) = C(\theta)[qI - A(\theta)]^{-1}B(\theta) \tag{4.79}$$

Although sampling a time-continuous description is a natural way to obtain the model (4.77), it could also for certain applications be posed directly in discrete time, with the matrices A, B, and C directly parametrized in terms of θ, rather than indirectly via (4.65).

Noise Representation and the Time-Invariant Kalman Filter

In the representation (4.77) and (4.78) we could further model the properties of the noise term $\{v(t)\}$. A straightforward but entirely valid approach would be to postulate a noise model of the kind

$$v(t) = H(q,\theta)e(t) \tag{4.80}$$

with $\{e(t)\}$ being white noise with variance λ. The θ-parameters in $H(q,\theta)$ could be partly in common with those in $G(q,\theta)$ or be extra additional noise model parameters.

For state-space descriptions, it is, however, more common to split the lumped noise term $v(t)$ into contributions from *measurement noise* $v(t)$ and *process noise* $w(t)$ acting on the states, so that (4.77) is written

$$\begin{aligned} x(t+1) &= A(\theta)x(t) + B(\theta)u(t) + w(t) \\ y(t) &= C(\theta)x(t) + v(t) \end{aligned} \tag{4.81}$$

Here $\{w(t)\}$ and $\{v(t)\}$ are assumed to be sequences of independent random variables with zero mean values and covariances

$$\begin{aligned} Ew(t)w^T(t) &= R_1(\theta) \\ Ev(t)v^T(t) &= R_2(\theta) \\ Ew(t)v^T(t) &= R_{12}(\theta) \end{aligned} \tag{4.82}$$

The disturbances $w(t)$ and $v(t)$ may often be signals whose physical origins are known. In Example 4.1 the load variation $T_\ell(t)$ was a "process noise," while the inaccuracy in the potentiometer angular sensor $v(t)$ was the "measurement noise." In such cases it may of course not always be realistic to assume that these signals are white noises. To arrive at (4.81) and (4.82) will then require extra modeling and extension of the state vector. See Problem 4G.2.

Let us now turn to the problem of predicting $y(t)$ in (4.81). This state-space description is one to which the celebrated Kalman filter applies (see, e.g., Anderson and Moore, 1979, for a thorough treatment). The conditional expectation of $y(t)$, given data $y(s)$, $u(s)$, $s \le t-1$ (i.e., from the infinite past up to time $t-1$), is, provided v and w are Gaussian processes, given by

$$\begin{aligned} \hat{x}(t+1,\theta) &= A(\theta)\hat{x}(t,\theta) + B(\theta)u(t) + K(\theta)[y(t) - C(\theta)\hat{x}(t,\theta)] \\ \hat{y}(t|\theta) &= C(\theta)\hat{x}(t,\theta) \end{aligned} \tag{4.83}$$

Here $K(\theta)$ is given as

$$K(\theta) = [A(\theta)\bar{P}(\theta)C^T(\theta) + R_{12}(\theta)][C(\theta)\bar{P}(\theta)C^T(\theta) + R_2(\theta)]^{-1} \quad (4.84a)$$

where $\bar{P}(\theta)$ is obtained as the positive semidefinite solution of the stationary Riccati equation:

$$\bar{P}(\theta) = A(\theta)\bar{P}(\theta)A^T(\theta) + R_1(\theta) - [A(\theta)\bar{P}(\theta)C^T(\theta) + R_{12}(\theta)][C(\theta)\bar{P}(\theta)C^T(\theta) + R_2(\theta)]^{-1}[A(\theta)\bar{P}(\theta)C^T(\theta) + R_{12}(\theta)]^T \quad (4.84b)$$

The predictor filter can thus be written as

$$\hat{y}(t|\theta) = C(\theta)[qI - A(\theta) + K(\theta)C(\theta)]^{-1}B(\theta)u(t) + C(\theta)[qI - A(\theta) + K(\theta)C(\theta)]^{-1}K(\theta)y(t) \quad (4.85)$$

The matrix $\bar{P}(\theta)$ is the covariance matrix of the state estimate error:

$$\bar{P}(\theta) = \bar{E}[x(t) - \hat{x}(t,\theta)][x(t) - \hat{x}(t,\theta)]^T \quad (4.86)$$

Innovations Representation

The prediction error

$$y(t) - C(\theta)\hat{x}(t,\theta) = C(\theta)[x(t) - \hat{x}(t,\theta)] + v(t) \quad (4.87)$$

in (4.83) amounts to that part of $y(t)$ that cannot be predicted from past data: "the innovation." Denoting this quantity by $e(t)$ as in (3.25), we find that (4.83) can be rewritten as

$$\begin{aligned} \hat{x}(t+1,\theta) &= A(\theta)\hat{x}(t,\theta) + B(\theta)u(t) + K(\theta)e(t) \\ y(t) &= C(\theta)\hat{x}(t,\theta) + e(t) \end{aligned} \quad (4.88a)$$

The covariance of $e(t)$ can be determined from (4.87) and (4.86):

$$Ee(t)e^T(t) = \Lambda = C(\theta)\bar{P}(\theta)C^T(\theta) + R_2(\theta) \quad (4.88b)$$

Since $e(t)$ appears explicitly, this representation is known as the *innovations form* of the state-space description. Using the shift operator q, we can clearly rearrange it as

$$y(t) = G(q,\theta)u(t) + H(q,\theta)e(t) \quad (4.89a)$$

$$\begin{aligned} G(q,\theta) &= C(\theta)[qI - A(\theta)]^{-1}B(\theta) \\ H(q,\theta) &= C(\theta)[qI - A(\theta)]^{-1}K(\theta) + I \end{aligned} \quad (4.89b)$$

showing its relationship to the general model (4.4) and to a direct modeling of $v(t)$ as in (4.80). See also Problem 4G.3.

In (4.88) the Kalman gain $K(\theta)$ is computed from $A(\theta)$, $C(\theta)$, $R_1(\theta)$, $R_{12}(\theta)$, and $R_2(\theta)$ in the fairly complicated manner given by (4.84). It is an attractive idea to sidestep (4.84) and the parametrization of the R-matrices by directly parametrizing $K(\theta)$ in terms of θ. This has the important advantage that the predictor (4.85) becomes a much simpler function of θ. Such a model structure we call a *directly parametrized innovations form.*

The R-matrices describing the noise properties contain $\frac{1}{2}n(n+1) + np + \frac{1}{2}p(p+1)$ matrix elements (discounting symmetric ones), while the Kalman gain K contains np elements ($p = \dim y$, $n = \dim x$). If we have no prior knowledge about the R-matrices and thus would need many parameters to describe them, it would therefore be a better alternative to parametrize $K(\theta)$, also from the point of view of keeping dim θ small. On the other hand, physical insight into (4.81) may entail knowing, for example, that the process noise affects only one state and is independent of the measurement noise, which might have a known variance. Then the parametrization of $K(\theta)$ via (4.82) and (4.84) may be done using less parameters than would be required in a direct parametrization of $K(\theta)$.

Remark. The parametrization in terms of (4.82) also gives a parametrization of the $p(p+1)/2$ elements of $\Lambda(\theta)$ in (4.88b). A direct parametrization of (4.88) would involve extra parameters for Λ, which, however, would not affect the predictor. (Compare also Problems 7E.4 and 8E.2.) ■

Directly parametrized innovations forms also contain black-box models that are in close relationship to those discussed in Section 4.2.

Example 4.2. Companion Form Parametrizations

In (4.88) let

$$\theta^T = [a_1\ a_2\ a_3\ b_1\ b_2\ b_3\ k_1\ k_2\ k_3]$$

and

$$A(\theta) = \begin{bmatrix} -a_1 & 1 & 0 \\ -a_2 & 0 & 1 \\ -a_3 & 0 & 0 \end{bmatrix}$$

$$B(\theta) = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \qquad K(\theta) = \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix}$$

$$C(\theta) = [1 \quad 0 \quad 0]$$

These matrices are said to be in *companion form* or in observer canonical form (see, e.g., Kailath, 1980). It is easy to verify that with these matrices

$$C(\theta)[qI - A(\theta)]^{-1}B(\theta) = \frac{b_1q^{-1} + b_2q^{-2} + b_3q^{-3}}{1 + a_1q^{-1} + a_2q^{-2} + a_3q^{-3}}$$

and

$$C(\theta)[qI - A(\theta)]^{-1}K(\theta) = \frac{k_1 q^{-1} + k_2 q^{-2} + k_3 q^{-3}}{1 + a_1 q^{-1} + a_2 q^{-2} + a_3 q^{-3}}$$

so that

$$1 + C(\theta)[qI - A(\theta)]^{-1}K(\theta) = \frac{1 + c_1 q^{-1} + c_2 q^{-2} + c_3 q^{-3}}{1 + a_1 q^{-1} + a_2 q^{-2} + a_3 q^{-3}}$$

with

$$c_i \triangleq a_i + k_i, \qquad i = 1, 2, 3$$

With this we have consequently obtained a parametrization of the ARMAX model set (4.15) and (4.16) for $n_a = n_b = n_c = 3$. ■

The corresponding parametrization of a multioutput model is more involved and is described in Appendix 4A.

Time-varying Predictors (*)

For the predictor filter (4.83) and (4.84) we assumed all previous data from time minus infinity to be available. If data prior to time $t = 0$ are lacking, we could replace them by zero, thus starting the recursion (4.83) at $t = 0$ with $\hat{x}(0) = 0$, and take the penalty of a suboptimal estimate. This was also our philosophy in Section 3.2.

An advantage with the state-space formulation is that a correct treatment of incomplete information about $t < 0$ can be given at the price of a slightly more complex predictor. If the information about the history of the system prior to $t = 0$ is given in terms of an initial state estimate $x_0(\theta) = \hat{x}(0,\theta)$ and associated uncertainty

$$\Pi_0(\theta) = E[x(0) - x_0(\theta)][x(0) - x_0(\theta)]^T \tag{4.90}$$

then the Kalman filter tells us that the one-step-ahead prediction is given by, (see, e.g., Anderson and Moore, 1979),

$$\hat{x}(t+1,\theta) = A(\theta)\hat{x}(t,\theta) + B(\theta)u(t) + K(t,\theta)[y(t) - C(\theta)\hat{x}(t,\theta)] \tag{4.91}$$

$$\hat{y}(t|\theta) = C(\theta)\hat{x}(t,\theta), \qquad \hat{x}(0,\theta) = x_0(\theta)$$

$$K(t,\theta) = [A(\theta)P(t,\theta)C^T(\theta) + R_{12}(\theta)][C(\theta)P(t,\theta)C^T(\theta) + R_2(\theta)]^{-1} \tag{4.92}$$

$$P(t+1,\theta) = A(\theta)P(t,\theta)A^T(\theta) + R_1(\theta) - K(t,\theta)[C(\theta)P(t,\theta)C^T(\theta) + R_2(\theta)]K^T(t,\theta), \qquad P(0,\theta) = \Pi_0(\theta)$$

Now $K(t,\theta)$ determined by (4.92) converges, under general conditions, fairly rapidly to $K(\theta)$ given by (4.84) (see, e.g., Anderson and Moore, 1979). For many problems it is thus reasonable to apply the limit form (4.83) with (4.84) directly to simplify calculations. For short data records, though, the solution (4.90) to (4.92) gives a useful possibility to deal with the transient properties in a correct way,

including possibly a parametrization of the unknown initial conditions $x_0(\theta)$ and $\Pi_0(\theta)$. Clearly, the steady-state approach (4.83) with (4.84) is a special case of (4.91) to (4.92), corresponding to $x_0(\theta) = 0$, $\Pi_0(\theta) = \bar{P}(\theta)$.

Sampling Continuous-time Process Noise (*)

Just as for the systems dynamics, we may have more insight into the nature of the process noise in continuous time. We could then pose a disturbed state-space model

$$\dot{x}(t) = F(\theta)x(t) + G(\theta)u(t) + \overline{w}(t) \tag{4.93}$$

where $\overline{w}(t)$ is formal white noise with covariance function

$$E\overline{w}(t)\overline{w}^T(s) = \overline{R}_1(\theta)\delta(t - s) \tag{4.94}$$

where δ is Dirac's delta function. When the input is piecewise constant as in (4.63), the corresponding discrete-time state equation becomes

$$x(kT + T) = A_T(\theta)x(kT) + B_T(\theta)u(kT) + w_T(kT) \tag{4.95}$$

where A_T and B_T are given by (4.65) and $w_T(kT)$, $k = 1, 2, \cdots$ is a sequence of independent random vectors with zero means and covariance matrix

$$Ew_T(kT)\, w_T^T(kT) = R_1(\theta) = \int_0^T e^{F(\theta)\tau}\, \overline{R}_1(\theta)\, e^{F^T(\theta)\tau} d\tau \tag{4.96}$$

See Åström (1970) for a derivation.

State-space Models

In summary, we have found that state-space models provide us with a spectrum of modeling possibilities: We may use physical modeling in continuous time with or without a corresponding time-continuous noise description to obtain structures with physical parameters θ. We can use physical parametrization of the dynamics part combined with a black-box parametrization of the noise properties, such as in the directly parametrized innovations form (4.88), or we can arrive at a noise model that is also physically parametrized via (4.93) to (4.96). Finally, we can use black-box state-space structures, such as the one of Example 4.2. These have the advantage over the input–output black box that the flexibility in choice of representation can secure better numerical properties of the parametrization (Problem 16E.1).

4.4 DISTRIBUTED PARAMETER MODELS (*)

Models that involve partial differential equations (PDE), directly or indirectly, when relating the input signal to the output signal are usually called *distributed parameter models*. "Distributed" then refers to the state vector, which in general

belongs to a function space, rather than $\mathbf{R}^n$. There are basically two ways to deal with such models. One is to replace the space variable derivative by a difference expression or to truncate a function series expansion so as to approximate the PDE by an ordinary differential equation. Then a "lumped" finite-dimensional model, of the kind we discussed in Section 4.3, is obtained. ("Lumped" refers to the fact that the distributed states are lumped together into a finite collection.) The other approach is to stick to the original PDE for the calculations, and only at the final, numerical, stage introduce approximations to facilitate the computations. It should be noted that this second approach also remains within the general model structure (4.4), provided the underlying PDE is linear and time invariant. This is best illustrated by an example.

Example 4.3 Heating Dynamics

Consider the physical system schematically depicted in Figure 4.7. It consists of a well-insulated metal rod, which is heated at one end. The heating effect at time t is the input $u(t)$, while the temperature measured at the other end is the output $y(t)$. This output is sampled at $t = 1, 2, \ldots$.

Under ideal conditions, this system is described by the heat-diffusion equation. If $x(t, \xi)$ denotes the temperature at time t, ξ length units from one end of the rod, then

$$\frac{\partial x(t, \xi)}{\partial t} = \kappa \frac{\partial^2 x(t, \xi)}{\partial \xi^2} \tag{4.97}$$

where κ is the coefficient of thermal conductivity. The heating at the far end means that

$$\left.\frac{\partial x(t, \xi)}{\partial \xi}\right|_{\xi = L} = K \cdot u(t) \tag{4.98}$$

where K is a heat-transfer coefficient. The near end is insulated so that

$$\left.\frac{\partial x(t, \xi)}{\partial \xi}\right|_{\xi = 0} = 0 \tag{4.99}$$

The measurements are

$$y(t) = x(t, 0) + v(t), \qquad t = 1, 2, \ldots \tag{4.100}$$

where $\{v(t)\}$ accounts for the measurement noise. The unknown parameters are

$$\theta = \begin{bmatrix} \kappa \\ K \end{bmatrix} \tag{4.101}$$

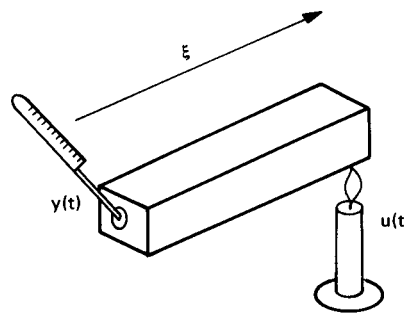

Figure 4.7 The heat-rod system.

Approximating

$$\frac{\partial^2 x(t,\xi)}{\partial \xi^2} = \frac{x(t,\xi+\Delta L) - 2x(t,\xi) + x(t,\xi-\Delta L)}{(\Delta L)^2}, \qquad \xi = k\cdot\Delta L$$

transfers (4.97) to a state-space model of order $n = \Delta L/L$, where the state variables $x(t, k\cdot\Delta L)$ are lumped representatives for $x(t,\xi), k\cdot\Delta L \le \xi < (k+1)\cdot\Delta L$. This often gives a reasonable approximation of the heat-diffusion equation.

Here we instead retain the PDE (4.97) by Laplace transforming it. Thus let $X(s,\xi)$ be the Laplace transform of $x(t,\xi)$ with respect to t for fixed ξ. Then (4.97) to (4.99) take the form

$$\begin{aligned} X(s,\xi) &= \kappa X''(s,\xi) \\ X'(s,L) &= K\cdot U(s) \\ X'(s,0) &= 0 \end{aligned} \tag{4.102}$$

Prime and double prime here denote differentiation with respect to ξ, and $U(s)$ is the Laplace transform of $u(t)$. Solving (4.102) for fixed s gives

$$X(s,\xi) = A(s)e^{-\xi\sqrt{s/\kappa}} + B(s)e^{\xi\sqrt{s/\kappa}}$$

where the constants $A(s)$ and $B(s)$ are determined from the boundary values

$$\begin{aligned} X'(s,0) &= 0 \\ X'(s,L) &= K\cdot U(s) \end{aligned}$$

which gives

$$A(s) = B(s) = \frac{K\cdot U(s)}{\sqrt{s/\kappa}(e^{L\sqrt{s/\kappa}} - e^{-L\sqrt{s/\kappa}})} \tag{4.103}$$

Inserting this into (4.100) gives

$$Y(s) = X(s,0) + V(s) = G_c(s,\theta)U(s) + V(s) \tag{4.104}$$

$$G_c(s,\theta) = \frac{2\kappa}{\sqrt{s/\kappa}(e^{L\sqrt{s/\kappa}} - e^{-L\sqrt{s/\kappa}})} \tag{4.105}$$

where $V(s)$ is the Laplace transform of the noise $\{v(t)\}$. We have thus arrived at a model parametrization of the kind (4.46). With some sampling procedure and a model for the measurement noise sequence, it can be carried further to the form (4.4). Note that $G_c(s,\theta)$ is an analytic function of s although not rational. All our concepts of poles, zeros, stability, and so on, can still be applied. ■

We can thus include distributed parameter models in our treatment of system identification methods. There is a substantial literature on this subject. See, for example, Banks, Crowley, and Kunisch (1983) and Kubrusly (1977). Not surprisingly, computational issues, choice of basis functions, and the like, play an important role in this literature.

4.5 MODEL SETS, MODEL STRUCTURES, AND IDENTIFIABILITY: SOME FORMAL ASPECTS (*)

In this chapter we have dealt with models of linear systems, as well as with parametrized sets of such models. When it comes to analysis of identification methods, it turns out that certain properties will have to be required from these models and model sets. In this section we shall discuss such formal aspects. To keep notation simple, we treat explicitly only SISO models.

Some Notation

For the expressions we shall deal with in this section, it is convenient to introduce some more compact notation. With

$$T(q) = [G(q)\ H(q)] \quad \text{and} \quad \chi(t) = \begin{bmatrix} u(t) \\ e(t) \end{bmatrix} \tag{4.106}$$

we can rewrite (4.1) as

$$y(t) = T(q)\chi(t) \tag{4.107}$$

The model structure (4.4) can similarly be written

$$y(t) = T(q,\theta)\chi(t), \qquad T(q,\theta) = [G(q,\theta)\ H(q,\theta)] \tag{4.108}$$

Given the model (4.107), we can determine the one-step-ahead predictor (3.56), which we can rewrite as

$$\hat{y}(t|t-1) = W(q)z(t) \tag{4.109}$$

with

$$W(q) = [W_u(q)\ W_y(q)], \qquad z(t) = \begin{bmatrix} u(t) \\ y(t) \end{bmatrix} \tag{4.110}$$

$$W_u(q) = H^{-1}(q)G(q), \qquad W_y(q) = [1 - H^{-1}(q)] \tag{4.111}$$

Clearly, (4.111) defines a one-to-one relationship between $T(q)$ and $W(q)$:

$$T(q) \leftrightarrow W(q) \tag{4.112}$$

Remark. Based on (4.107), we may prefer to work with the k-step-ahead predictor (3.31). To keep the link (4.112), we can view (3.31) as the one-step-ahead predictor for the model (3.32). ■

Models

We noted already in (4.1) that a model of a linear system consists of specified transfer functions $G(z)$ and $H(z)$, possibly complemented with a specification of the prediction error variance λ, or the PDF $f_e(x)$ of the prediction error e. In Sections 3.2 and 3.3, we made the point that what matters in the end is by which expression

future outputs are predicted. The one-step-ahead predictor based on the model (4.1) is given by (4.109).

While the predictor (4.109) via (4.112) is in a one-to-one relationship with (4.107), it is useful to relax the link (4.112) and regard (4.109) as the basic model. This will, among other things, allow a direct extension to nonlinear and time-varying models, as shown in Section 5.4. We may thus formally define what we mean by a model:

Definition 4.1. A *predictor model* of a linear, time-invariant system is a stable filter $W(q)$, defining a predictor (4.109) as in (4.110). ■

Stability, which was defined in (2.27) [applying to both components of $W(q)$] is necessary to make the right side of (4.109) well defined. While predictor models are meaningful also in a deterministic framework without a stochastic alibi, as discussed in Section 3.3, it is useful also to consider models that specify properties of the associated prediction errors (innovations).

Definition 4.2. A *complete probabilistic model* of a linear, time-invariant system is a pair $(W(q), f_e(x))$ of a predictor model $W(q)$ and the PDF $f_e(x)$ of the associated prediction errors. ■

Clearly, we can also have models where the PDFs are only partially specified (e.g., by the variance of e).

In this section we shall henceforth only deal with predictor models and therefore drop the adjective. The concepts for probabilistic models are quite analogous.

We shall say that two models $W_1(q)$ and $W_2(q)$ are *equal* if

$$W_1(e^{i\omega}) = W_2(e^{i\omega}), \qquad \text{almost all } \omega \tag{4.113}$$

A model

$$W(q) = [W_u(q)\ W_y(q)]$$

will be called a *k-step-ahead predictor model* if

$$W_y(q) = \sum_{\ell=k}^{\infty} w_y(\ell) q^{-\ell}, \qquad \text{with } w_y(k) \neq 0 \tag{4.114}$$

and an *output error model* (or a *simulation model*) if $W_y(q) \equiv 0$.

Note that the definition requires the predictors to be stable. This does not necessarily mean that the system dynamics is stable.

Example 4.4 Unstable System

Suppose that

$$G(q) = \frac{bq^{-1}}{1 + aq^{-1}}, \qquad \text{with } |a| > 1$$

and

$$H(q) = \frac{1}{1 + aq^{-1}}$$

This means that the model is described by

$$y(t) + ay(t-1) = bu(t-1) + e(t)$$

and the dynamics from u to y is unstable. The predictor functions are, however:

$$W_y(q) = -aq^{-1}, \qquad W_u(q) = bq^{-1}$$

implying that

$$\hat{y}(t|t-1) = -ay(t-1) + bu(t-1)$$

which clearly satisfies the condition of Definition 4.1. ■

Model Sets

Definition 4.1 describes one given model of a linear system. The identification problem is to determine such a model. The search for a suitable model will typically be conducted over a set of candidate models. Quite naturally, we define a *model set* $\mathcal{M}^*$ as

$$\mathcal{M}^* = \{W_\alpha(q) | \alpha \in \mathcal{A}\} \tag{4.115}$$

This is just a collection of models, each subject to Definition 4.1, here "enumerated" with an index α covering an index set $\mathcal{A}$.

Typical model sets could be

$$\mathcal{M}^* = \mathcal{L}^* = \{\text{all linear models}\}$$

that is, all models that are subject to Definition 4.1, or

$$\mathcal{M}_n^* = \{\text{all models such that } W_y(q) \text{ and } W_u(q) \text{ are polynomials of } q^{-1} \text{ of degree at most } n\} \tag{4.116}$$

or a finite model set

$$\mathcal{M}^* = \{W_1(q), W_2(q), W_3(q)\} \tag{4.117}$$

We say that *two model sets are equal*, $\mathcal{M}_1^* = \mathcal{M}_2^*$, if for any W_1 in $\mathcal{M}_1^*$ there exists a W_2 in $\mathcal{M}_2^*$ such that $W_1 = W_2$ [defined by (4.113)], and vice versa.

Model Structures: Parametrization of Model Sets

Most often a model set of interest is noncountable. Since we have to conduct a search over it for "the best model," it is then interesting how the indexation is chosen. The basic idea is to parametrize (index) the set "smoothly" over a "nice"

area and perform the search over the parameter set (the index set). To put this formally, we let the model be indexed by a d-dimensional vector θ:

$$W(q,\theta)$$

To formalize "smoothly," we require that for any given z, $|z| \geq 1$, the complex-valued function $W(z,\theta)$ of θ be differentiable:

$$\Psi(z,\theta) = \frac{d}{d\theta} W(z,\theta) \tag{4.118a}$$

Here

$$\Psi(z,\theta) = [\Psi_u(z,\theta)\ \Psi_y(z,\theta)] = \left[\frac{d}{d\theta} W_u(z,\theta)\ \frac{d}{d\theta} W_y(z,\theta)\right] \tag{4.118b}$$

is a $d \times 2$ matrix. Thus the gradient of the prediction $\hat{y}(t|\theta)$ is given by

$$\psi(t,\theta) = \frac{d}{d\theta}\hat{y}(t|\theta) = \Psi(q,\theta)z(t) \tag{4.118c}$$

Since the filters Ψ will have to be computed and used when the search is carried out, we also require them to be stable. We thus have the following definition:

Definition 4.3. *A model structure* $\mathcal{M}$ is a differentiable mapping from a connected, open subset $D_{\mathcal{M}}$ of $\mathbf{R}^d$ to a model set $\mathcal{M}^*$, such that the gradients of the predictor functions are stable. ■

To put this definition in mathematical notation we have

$$\mathcal{M}\colon D_{\mathcal{M}} \ni \theta \to \mathcal{M}(\theta) = W(q,\theta) \in \mathcal{M}^* \tag{4.119}$$

such that the filter Ψ in (4.118) exists and is stable for $\theta \in D_{\mathcal{M}}$. We will thus use $\mathcal{M}(\theta)$ to denote the particular model corresponding to θ and reserve $\mathcal{M}$ for the mapping itself.

Remark. The requirement that $D_{\mathcal{M}}$ should be open is in order for the derivatives in (4.118) to be unambiguously well defined. When using model structures, we may sometimes prefer to work with nonopen sets $D_{\mathcal{M}}$. Clearly, as long as $D_{\mathcal{M}}$ is contained in an open set where (4.118) are defined, no problems will occur. Differentiability can also be defined over more complicated subsets of $\mathbf{R}^d$ than open ones, that is, *differentiable manifolds* (see, e.g., Boothby, 1975). See the chapter bibliography for further comments. ■

Example 4.5 An ARX Structure

Consider the ARX model

$$y(t) + ay(t-1) = b_1 u(t-1) + b_2 u(t-2) + e(t)$$

The predictor is given by (4.10), which means that

$$W(q,\theta) = [b_1 q^{-1} + b_2 q^{-2}\ \ -aq^{-1}], \qquad \theta = [a\ b_1\ b_2]^T$$

and

$$\Psi(q,\theta) = \begin{bmatrix} 0 & -q^{-1} \\ q^{-1} & 0 \\ q^{-2} & 0 \end{bmatrix} \quad \blacksquare$$

The parametrized model sets that we have explicitly studied in this chapter have been in terms of (4.4), that is,

$$y(t) = G(q,\theta)u(t) + H(q,\theta)e(t), \qquad \theta \in D_{\mathcal{M}} \tag{4.120}$$

or using (4.108)

$$y(t) = T(q,\theta)\chi(t)$$

It is immediate to verify that, in view of (4.111),

$$\Psi(q,\theta) = \frac{1}{(H(q,\theta))^2}\mathbf{T}'(q,\theta)\begin{bmatrix} H(q,\theta) & 0 \\ -G(q,\theta) & 1 \end{bmatrix} \tag{4.121}$$

where $\mathbf{T}'(q,\theta)$ is the $d \times 2$ matrix

$$\mathbf{T}'(q,\theta) = \frac{d}{d\theta}T(q,\theta) = \left[\frac{d}{d\theta}G(q,\theta)\ \frac{d}{d\theta}H(q,\theta)\right] \tag{4.122}$$

Differentiability of W is thus assured by differentiability of T.

It should be clear that all parametrizations we have considered in this chapter indeed are model structures in the sense of Definition 4.3. We have, for example:

Lemma 4.1. The parametrization (4.35) together with (4.41) with θ confined to $D_{\mathcal{M}} = \{\theta | F(z) \cdot C(z)$ has no zeros on or outside the unit circle$\}$ is a model structure.

Proof. We need only verify that the gradients of

$$W_u(z,\theta) = \frac{B(z)D(z)}{C(z)F(z)}$$

and

$$W_y(z,\theta) = 1 - \frac{D(z)A(z)}{C(z)}$$

with respect to θ are analytical in $|z| \geq 1$ for $\theta \in D_{\mathcal{M}}$. But this is immediate since, for example

$$\frac{\partial}{\partial c_k}W_u(z,\theta) = -\frac{B(z)D(z)z^{-k}}{[C(z)]^2F(z)} \quad \blacksquare$$

Lemma 4.2. Consider the state-space parametrization (4.88). Assume that the entries of the matrices $A(\theta)$, $B(\theta)$, $K(\theta)$, and $C(\theta)$ are differentiable with respect to θ. Suppose that $\theta \in D_{\mathcal{M}}$, with

$$D_{\mathcal{M}} = \{\theta | \text{all eigenvalues of } A(\theta) - K(\theta)C(\theta) \text{ are inside the unit circle}\}$$

Then the parametrization of the corresponding predictor is a model structure.

Proof. See Problem 4D.1. ■

Notice that when $K(\theta)$ is obtained as the solution of (4.84), then by a standard Kalman filter property (see Anderson and Moore, 1979),

$$D_{\mathcal{M}} = \{\theta | [A(\theta), R_1(\theta)] \text{ stabilizable and } [A(\theta), C(\theta)] \text{ detectable}\} \quad (4.123)$$

When relating different model structures, we shall use the following concept.

Definition 4.4. A model structure $\mathcal{M}_1$ is said to be *contained in* $\mathcal{M}_2$,

$$\mathcal{M}_1 \subset \mathcal{M}_2 \quad (4.124)$$

if $D_{\mathcal{M}_1} \subset D_{\mathcal{M}_2}$ and the mapping $\mathcal{M}_1$ is obtained by restricting $\mathcal{M}_2$ to $\theta \in D_{\mathcal{M}_1}$. The archetypical situation for (4.124) is when $\mathcal{M}_2$ defines nth-order models and $\mathcal{M}_1$ defines mth-order models, $m < n$. One could think of $\mathcal{M}_1$ as obtained from $\mathcal{M}_2$ by fixing some parameters (typically to zero). ■

The following property of a model structure is sometimes useful:

Definition 4.5. A model structure $\mathcal{M}$ is said to have an *independently parametrized transfer function and noise model* if

$$\theta = \begin{bmatrix} \rho \\ \eta \end{bmatrix}, \qquad D_{\mathcal{M}} = D_\rho \times D_\eta, \qquad \rho \in D_\rho, \qquad \eta \in D_\eta$$
$$T(q, \theta) = [G(q, \rho)\ H(q, \eta)] \quad \blacksquare \quad (4.125)$$

We note that in the family (4.33) the special cases with $A(q) \equiv 1$ correspond to independent parametrizations of G and H.

Remark On "Finite Model Structures": Sometimes the set of candidate models is finite as in (4.117). It may still be desirable to index it using a parameter vector θ, now ranging over a finite set of points. Although such a construction does not qualify as a "model structure" according to Definition 4.3, it should be noted that the estimation procedures of Sections 7.1 to 7.4, as well as the convergence analysis of Sections 8.1 to 8.5, still make sense in this case. ■

Model Set as a Range of a Model Structure

A model structure will clearly define a model set by its range:

$$\mathcal{M}^* = \mathcal{R}(\mathcal{M}) = \text{Range } \mathcal{M} = \{\mathcal{M}(\theta) | \theta \in D_{\mathcal{M}}\}$$

An important problem for system identification is to find a model structure whose range equals a given model set. This may sometimes be an easy problem and sometimes highly nontrivial.

Example 4.6 Parametrizing $\mathcal{M}_3^*$

Consider the set $\mathcal{M}_n^*$ defined by (4.116) with $n = 3$. If we take

$$\theta = [a_1\ a_2\ a_3\ b_1\ b_2\ b_3]^T, \qquad d = 6$$

$$D_{\mathcal{M}} = \mathbf{R}^6$$

and

$$W_y(q, \theta) = -a_1 q^{-1} - a_2 q^{-2} - a_3 q^{-3}$$

$$W_u(q, \theta) = b_1 q^{-1} + b_2 q^{-2} + b_3 q^{-3}$$

we have obviously constructed a model structure whose range equals $\mathcal{M}_3^*$. ■

A given model set can typically be described as the range of several different model structures (see Problems 4E.6 and 4E.9).

Model Set as a Union of Ranges of Model Structures

In the preceding example it was possible to describe the desired model set as the range of a model structure. We shall later encounter model sets for which this is not possible, at least not with model structures with desired identifiability properties. The remedy for these problems is to describe the model set as a union of ranges of different model structures:

$$\mathcal{M}^* = \bigcup_{i=1}^{\beta} \mathcal{R}(\mathcal{M}_i) \tag{4.126}$$

This idea has been pursued in particular for representing linear multioutput systems. We shall give the details of this procedure in Appendix 4A. Let us here only remark that model sets described by (4.126) are useful also for working with models of different orders, and that they are often used, at least implicitly, when the order of a suitable model is unknown and is to be determined.

Identifiability Properties

Identifiability is a concept that is central in identification problems. Loosely speaking, the problem is whether the identification procedure will yield a unique value of the parameter θ, and/or whether the resulting model is equal to the true system. We shall deal with the subject in more detail in the analysis chapter (see Sections 8.2 and 8.3). The issue involves aspects on whether the data set (the experimental conditions) is informative enough to distinguish between different models as well as properties of the model structure itself: If the data are informative enough to distinguish between nonequal models, then the question is whether different values of θ can give equal models. With our terminology, the latter prob-

lem concerns the *invertibility of the model structure* $\mathcal{M}$ (i.e., whether $\mathcal{M}$ is injective). We shall now discuss some concepts related to such invertibility properties. Remember that these are only one leg of the identifiability concept. They are to be complemented in Sections 8.2 and 8.3.

Definition 4.6. A model structure $\mathcal{M}$ is *globally identifiable at* θ^* if

$$\mathcal{M}(\theta) = \mathcal{M}(\theta^*), \quad \theta \in D_{\mathcal{M}} \Rightarrow \theta = \theta^* \tag{4.127}$$

Recall that model equality was defined in (4.113), requiring the predictor transfer functions to coincide. According to (4.112), this means that the underlying transfer functions G and H coincide. ■

Once identifiability at a point is defined, we proceed to properties of the whole set.

Definition 4.7. A model structure $\mathcal{M}$ is *strictly globally identifiable* if it is globally identifiable at all $\theta^* \in D_{\mathcal{M}}$. ■

This definition is quite demanding. As we shall see, it is difficult to construct model structures that are strictly globally identifiable. The difficulty for linear systems, for example, is that global identifiability may be lost at points on hypersurfaces corresponding to lower-order systems. Therefore, we introduce a weaker and more realistic property:

Definition 4.8. A model structure $\mathcal{M}$ is *globally identifiable* if it is globally identifiable at almost all $\theta^* \in D_{\mathcal{M}}$. ■

Remark. This means that $\mathcal{M}$ is globally identifiable at all $\theta^* \in \tilde{D}_{\mathcal{M}} \subset D_{\mathcal{M}}$, where

$$\delta D_{\mathcal{M}} = \{\theta | \theta \in D_{\mathcal{M}}; \theta \notin \tilde{D}_{\mathcal{M}}\}$$

is a set of Lebesgue measure zero in $\mathbf{R}^d$ (recall that $D_{\mathcal{M}}$ and hence $\delta D_{\mathcal{M}}$ is a subset of $\mathbf{R}^d$). ■

For corresponding local properties, the most natural definition of local identifiability of $\mathcal{M}$ at θ^* would be to require that there exists an ε such that

$$\mathcal{M}(\theta) = \mathcal{M}(\theta^*), \quad \theta \in \mathcal{B}(\theta^*, \varepsilon) \Rightarrow \theta = \theta^* \tag{4.128}$$

where $\mathcal{B}(\theta^*, \varepsilon)$ denotes an ε-neighborhood of θ^*.

(Strict) local identifiability of a model structure can then be defined analogously to Definitions 4.7 and 4.8. See also Problem 4G.4.

Use of the Identifiability Concept

The identifiability concept concerns the unique representation of a given system description in a model structure. Let

$$\mathcal{S}: y(t) = G_0(q)u(t) + H_0(q)e(t) \tag{4.129}$$

be such a description. We could think of it as a "true" or "ideal" description of the actual system, but such an interpretation is immaterial for the moment. Let $\mathcal{M}$ be a model structure based on one-step-ahead predictors for

$$y(t) = G(q,\theta)u(t) + H(q,\theta)e(t) \tag{4.130}$$

Then define the set $D_T(\mathcal{S},\mathcal{M})$ as those θ-values in $D_{\mathcal{M}}$ for which $\mathcal{S} = \mathcal{M}(\theta)$. We can write this as

$$D_T(\mathcal{S},\mathcal{M}) = \{\theta \in D_{\mathcal{M}} | G_0(z) = G(z,\theta),\ H_0(z) = H(z,\theta) \text{ almost all } z\} \tag{4.131}$$

This set is empty in case $\mathcal{S} \notin \mathcal{M}$. (Here, with abuse of notation, $\mathcal{M}$ also denotes the range of the mapping $\mathcal{M}$.)

Now suppose that $\mathcal{S} \in \mathcal{M}$ so that $\mathcal{S} = \mathcal{M}(\theta_0)$ for some value θ_0. Furthermore, suppose that $\mathcal{M}$ *is globally identifiable at* θ_0. Then

$$D_T(\mathcal{S},\mathcal{M}) = \{\theta_0\} \tag{4.132}$$

One aspect of the choice of a good model structure is to select $\mathcal{M}$ so that (4.132) holds for the given description $\mathcal{S}$. Since $\mathcal{S}$ is unknown to the user, this will typically involve tests of several different structures $\mathcal{M}$. The identifiability concepts will then provide useful guidance in finding an $\mathcal{M}$ such that (4.132) holds.

4.6 IDENTIFIABILITY OF SOME MODEL STRUCTURES

Definition 4.6 and (4.113) together imply that a model structure is globally identifiable at θ^* if and only if

$$G(z,\theta) \equiv G(z,\theta^*) \quad \text{and} \quad H(z,\theta) \equiv H(z,\theta^*) \quad \text{for almost all } z \Rightarrow \theta = \theta^* \tag{4.133}$$

For local identifiability, we consider only θ confined to a sufficiently small neighborhood of θ^*. A general approach to test local identifiability is given by the criterion in Problem 4G.4.

Global identifiability is more difficult to deal with in general terms. In this section we shall only briefly discuss identifiability of physical parameters and give some results for general black-box SISO models. Black-box multivariable systems are dealt with in Appendix 4A.

Parametrizations in Terms of Physical Parameters

Modeling physical processes typically leads to a continuous-time state-space model (4.59) to (4.60), summarized as (4.62) ($T = 1$):

$$y(t) = G_c(p,\theta)u(t) + v(t) \tag{4.134}$$

For proper handling we should sample G_c and include a noise model H so that (4.133) can be applied for identifiability tests. A simpler test to apply is

$$G_c(s,\theta) = G_c(s,\theta^*) \text{ almost all } s \Rightarrow \theta = \theta^*? \tag{4.135}$$

It is true that this is not identical to (4.133): When sampling G_c, ambiguities may occur; two different G_c can give the same G_T [cf. (2.24)]. Equation (4.135) is thus not sufficient for (4.133) to hold. However, with a carefully selected sampling interval, this ambiguity should not cause any problems. Also, a θ-parametrized noise model may help in resolving (4.135). This condition is thus not necessary for (4.133) to hold. However, in most applications the noise characteristics are not so significant that they indeed bear information about the physical parameters. All this means that (4.135) is a reasonable test for global identifiability of the corresponding model structure at θ^*.

Now (4.135) is a difficult enough problem. Except for special structures there are no general techniques available other than brute-force solution of the equations underlying (4.135). See Problems 4E.5 and 4E.6 for some examples. A comprehensive treatment of (4.135) for state-space models is given by Walter (1982), and Godfrey (1983) discusses the same problem for compartmental models. See also Godfrey and Distefano (1985).

SISO Transfer-function Model Structures

We shall now aim at an analysis of the general black-box SISO model structure (4.33) together with (4.41). Let us first illustrate the character of the analysis with two simple special cases.

Consider the ARX model structure (4.7) together with (4.9):

$$G(z,\theta) = \frac{B(z)}{A(z)}, \qquad H(z,\theta) = \frac{1}{A(z)} \tag{4.136}$$
$$\theta = [a_1 \dots a_{n_a} b_1 \dots b_{n_b}]^T$$

Equality for H in (4.133) implies that the A-polynomials coincide, which in turn implies that the B-polynomials must coincide for the G to be equal. It is thus immediate to verify that (4.133) holds for all θ^* in the model structure (4.136). Consequently, *the structure (4.136) is strictly globally identifiable.*

Let us now turn to the OE model structure (4.25) with orders n_b and n_f. At $\theta = \theta^*$ we have

$$\begin{aligned} G(z,\theta^*) &= \frac{B^*(z)}{F^*(z)} = \frac{b_1^* z^{-1} + \cdots + b_{n_b}^* z^{-n_b}}{1 + f_1^* z^{-1} + \cdots + f_{n_f}^* z^{-n_f}} \\ &= z^{n_f - n_b} \frac{b_1^* z^{n_b - 1} + \cdots + b_{n_b}^*}{z^{n_f} + f_1^* z^{n_f - 1} + \cdots + f_{n_f}^*} = z^{n_f - n_b} \frac{z^{n_b} B^*(z)}{z^{n_f} F^*(z)} \end{aligned} \tag{4.137}$$

We shall work with the polynomial $\tilde{F}^*(z) = z^{n_f} F^*(z)$ in the variable z, rather than with $F^*(z)$, which is a polynomial in z^{-1}. The reason is that $z^{n_f} F^*(z)$ always has

degree n_f regardless of whether $f^*_{n_f}$ is zero. Let $\tilde{B}^*(z) = z^{n_b} B^*(z)$, and let θ be an arbitrary parameter value. We can then write (4.133),

$$G(z,\theta^*) = G(z,\theta) = \frac{B(z)}{F(z)} = z^{n_f - n_b}\frac{\tilde{B}(z)}{\tilde{F}(z)}$$

as

$$\tilde{F}(z)\tilde{B}^*(z) - \tilde{F}^*(z)\tilde{B}(z) \equiv 0 \tag{4.138}$$

Since $\tilde{F}^*(z)$ is a polynomial of degree n_f, it has n_f zeros:

$$\tilde{F}^*(\alpha_i) = 0, \qquad i = 1, \ldots, n_f$$

Suppose that $\tilde{B}^*(\alpha_i) \neq 0, i = 1, \ldots, n_f$; that is, $\tilde{B}^*(z)$ and $\tilde{F}^*(z)$ are *coprime* (have no common factors). Then (4.138) implies that

$$\tilde{F}^*(\alpha_i) = 0, \qquad i = 1, \ldots, n_f$$

[if a zero α_i has multiplicity n_i, then differentiate (4.138) $n_i - 1$ times to conclude that it is a zero of the same multiplicity to $\tilde{F}(z)$]. Consequently, we have $\tilde{F}(z) \equiv \tilde{F}^*(z)$, which in turn implies that $\tilde{B}(z) = \tilde{B}^*(z)$ so that $\theta = \theta^*$. If, on the other hand, $\tilde{F}^*$ and $\tilde{B}^*$ do have a common factor so that

$$\tilde{F}^*(z) = \gamma(z)\tilde{F}^*_1(z), \qquad \tilde{B}^*(z) = \gamma(z)\tilde{B}^*_1(z)$$

then all θ, such that

$$\tilde{F}(z) = \beta(z)\tilde{F}^*_1(z), \qquad \tilde{B}(z) = \beta(z)\tilde{B}^*_1(z)$$

for arbitrary $\beta(z)$ will yield equality in (4.138). Hence the model structure is neither globally nor locally identifiable at θ^* [$\beta(z)$ can be chosen arbitrarily close to $\gamma(z)$]. We thus find that *the OE structure (4.25) is globally and locally identifiable at θ^* if and only if the corresponding numerator and denominator polynomials $z^{n_f}F^*(z)$ and $z^{n_b}B^*(z)$ are coprime.*

The generalization to the black-box SISO structure (4.33) is now straightforward:

Theorem 4.1. Consider the model structure $\mathcal{M}$ corresponding to

$$A(q)y(t) = \frac{B(q)}{F(q)}u(t) + \frac{C(q)}{D(q)}e(t) \tag{4.139}$$

with θ, given by (4.41), being the coefficients of the polynomials involved. The degrees of the polynomials are n_a, n_b, and so on. This model structure is locally and globally identifiable at θ^* if and only if all of (i) to (iv) hold:

(i) There is no common factor to all of $z^{n_a}A^*(z)$, $z^{n_b}B^*(z)$, and $z^{n_c}C^*(z)$.
(ii) There is no common factor to $z^{n_b}B^*(z)$ and $z^{n_f}F^*(z)$.
(iii) There is no common factor to $z^{n_c}C^*(z)$ and $z^{n_d}D^*(z)$.

(iv) If $n_a \geq 1$, then also require that there is no common factor to $z^{n_f}F^*(z)$ and $z^{n_d}D^*(z)$.

The starred polynomials correspond to θ^*. ■

Notice that several of the conditions (i) to (iv) will be automatically satisfied in the common special cases of (4.139). For example, when $C \equiv F \equiv 1$ [the case (4.22)], the model structure is globally identifiable at all θ^*. Notice also that any of the conditions (i) to (iv) can be violated only for "special" θ^*, placed on hypersurfaces in $\mathbf{R}^d$. We thus have the following corollary:

Corollary: The model structure given by (4.139) is globally identifiable. When the C and F polynomials are constrained to be of zero order, the corresponding model structure is strictly globally identifiable. ■

Looking for a "True" System Within Identifiable Structures

We shall now illustrate the usefulness of Theorem 4.1 by applying it to the problem of finding an $\mathcal{M}$ such that (4.132) holds for a given $\mathcal{S}$. Suppose that $\mathcal{S}$ is given by

$$\mathcal{S}\colon G_0(q) = \frac{B_0(q)}{A_0(q)F_0(q)}, \qquad H_0(q) = \frac{C_0(q)}{A_0(q)D_0(q)} \tag{4.140}$$

with orders n_a^0, n_b^0, and so on (after all possible cancellations of common factors). This system belongs to the model structure $\mathcal{M}$ in (4.139) provided all the model orders are at least as large as the true ones:

$$n_a \geq n_a^0, \qquad n_b \geq n_b^0, \text{ etc.} \tag{4.141}$$

When (4.141) holds, let θ_0 be a value that gives the description (4.140):

$$\mathcal{S} = \mathcal{M}(\theta_0) \tag{4.142}$$

Now, clearly, $\mathcal{M}$ will be globally identifiable at θ_0 and (4.132) will hold if we have equality in all of (4.141). The true orders $n_a^0, \ldots,$ are, however, typically not known, and it would be quite laborious to search for all combinations of model orders until equalities in (4.141) were obtained. The point of Theorem 4.1 is that such a search is not necessary; the structure $\mathcal{M}$ is globally identifiable at θ_0 under weaker conditions.

We have the following reformulation of Theorem 4.1:

Theorem 4.2. Consider the system description $\mathcal{S}$ in (4.140) with true polynomial orders n_a^0, n_b^0, and so on, as defined in the text. Consider model structure $\mathcal{M}$ of Theorem 4.1. Then $\mathcal{S} \in \mathcal{M}$ and corresponds to a globally identifiable θ-value if and only if

(i) $\min(n_a - n_a^0, n_b - n_b^0, n_c - n_c^0) = 0.$
(ii) $\min(n_b - n_b^0, n_f - n_f^0) = 0.$
(iii) $\min(n_c - n_c^0, n_d - n_d^0) = 0.$
(iv) If $n_a > 0$, then also $\min(n_f - n_f^0, n_d - n_d^0) = 0.$ ■

With Theorem 4.2, the search for a true system within identifiable model structures is simplified. If, for example, $\mathcal{S}$ can be described in ARMAX form with finite orders n_a^0, n_b^0, and n_c^0, then we may take $n_a = n_b = n_c = n$ ($n_f = n_d = 0$) in $\mathcal{M}$, giving a model structure, say, $\mathcal{M}_n$. By increasing n one unit at a time, we will sooner or later strike a structure where (i) holds and thus $\mathcal{S}$ can be uniquely represented.

SISO State-space Models

Consider now a state-space model structure (4.88). It is quite clear that the matrices $A(\theta)$, $B(\theta)$, $C(\theta)$, and $K(\theta)$ cannot be "filled" with parameters, since the corresponding input–output description (4.89) is defined by $3n$ parameters only ($n = \dim x$). To obtain identifiable structures, it is thus natural to seek parametrizations of the matrices that involve $3n$ parameters; the coefficients of the two $(n-1)$th order numerator polynomials and the coefficients of the common, monic nth order denominator polynomial or some transformation of these coefficients. One such parametrization is the observer canonical form of Example 4.2, which we can write in symbolic form as

$$x(t+1,\theta) = A(\theta)x(t,\theta) + B(\theta)u(t) + K(\theta)e(t)$$
$$y(t) = C(\theta)x(t,\theta) + e(t) \tag{4.143a}$$

$$A(\theta) = \begin{bmatrix} \times & & \\ \times & & I_{n-1} \\ \vdots & & \\ \times & 0 \ldots 0 & \end{bmatrix} \qquad B(\theta) = \begin{bmatrix} \times \\ \times \\ \vdots \\ \times \end{bmatrix}, \qquad K(\theta) = \begin{bmatrix} \times \\ \times \\ \vdots \\ \times \end{bmatrix} \tag{4.143b}$$

$$C(\theta) = [1 \quad 0 \ldots 0]$$

Here I_{n-1} is the $(n-1) \times (n-1)$ unit matrix, while $\times$ marks an adjustable parameter. This representation is observable by construction.

According to Example 4.2, this structure is in one-to-one correspondence with an ARMAX structure with $n_a = n_b = n_c = n$. From Theorem 4.1 we know that this is identifiable at θ^*, provided the corresponding polynomials do not all have a common factor, meaning that the model could be represented using a smaller value of n. It is well known that for state-space models this can only happen if the model is uncontrollable and/or unobservable. Since (4.143) is observable by construction, we thus conclude that *this structure is globally and locally identifiable at θ^* if and only if the two-input system $\{A(\theta^*), [B(\theta^*)\ K(\theta^*)]\}$ is controllable.* Note that this result applies to the particular state-space structure (4.143) only.

4.7 SUMMARY

In this chapter we have studied sets of predictors of the type

$$\hat{y}(t|\theta) = W_u(q,\theta)u(t) + W_y(q,\theta)y(t), \qquad \theta \in D_{\mathcal{M}} \subset \mathbf{R}^d \tag{4.144}$$

These are in one-to-one correspondence with model descriptions

$$y(t) = G(q,\theta)u(t) + H(q,\theta)e(t), \qquad \theta \in D_{\mathcal{M}} \tag{4.145}$$

with $\{e(t)\}$ as white noise, via

$$W_u(q,\theta) = H^{-1}(q,\theta)G(q,\theta)$$
$$W_y(q,\theta) = [1 - H^{-1}(q,\theta)]$$

When choosing models it is usually most convenient to go via (4.145), even if (4.144) is the "operational" version.

We have denoted parametrized model sets, or *model structures* by $\mathcal{M}$, while a particular model corresponding to the parameter value θ is denoted by $\mathcal{M}(\theta)$. Such a parametrization is instrumental in conducting a search for "best models." Two different philosophies may guide the choice of parametrized model sets:

1. *Black-box model structures:* The prime idea is to obtain flexible model sets that can accommodate a variety of systems, without looking into their internal structures. The input–output model structures of Section 4.2, as well as canonically parametrized state-space models (see Example 4.2), are of this character.
2. *Model structures with physical parameters:* The idea is to incorporate physical insight into the model set so as to bring the number of adjustable parameters down to what is actually unknown about the system. Continuous-time state-space models are typical representatives for this approach.

We have also in this chapter introduced formal requirements on the predictor filters $W_u(q,\theta)$ and $W_y(q,\theta)$ (Definition 4.3) and discussed concepts of parameter identifiability (i.e., whether the parameter θ can be uniquely determined from the predictor filters). These properties were investigated for the most typical black-box model structures in Section 4.6 and Appendix 4A. The bottom line of these results is that identifiability can be secured, provided certain orders are chosen properly. The number of such orders to be chosen typically equals the number of outputs.

4.8 BIBLIOGRAPHY

The selection of a parametrized set of models is, as we have noted, vital for the identification problem. This is the link between system identification and parameter estimation techniques. Most articles and books on system identification thus contain material on model structures, even if not presented in as explicit terms as here.

The simple equation error model (4.7) has been widely studied in many contexts. See, for example, Åström (1968), Hsia (1977), Mendel (1973), and Unbehauen, Göhring, and Bauer (1974) for discussions related to identification. Linear models like (4.12) are a prime object of study in statistics; see, for example, Rao (1973) or Draper and Smith (1981). The ARMAX model was introduced into system identification in Åström and Bohlin (1965) and is since then a basic model. The ARARX model structure was introduced into the control literature by Clarke (1967), but was apparently first used in a statistical framework by Cochrane and Orcutt (1949). The term pseudolinear regression for the representation (4.21) was introduced by Solo (1978). Output error models are treated, for example, in Dugard and Landau (1980) and Kabaila (1983). The general family (4.33) was first discussed in Ljung (1979b). It was used in Ljung and Söderström (1983). Multivariable MFDs are discussed in Kailath (1980) and Dickinson, Kailath, and Morf (1974). Corresponding right MFDs to (4.56), where the inverses are on the right, are discussed in Nehorai and Morf (1984). When no input is present, the corresponding model structures reduce to AR, MA, and ARMA descriptions. These are discussed in many textbooks on time series (e.g., Box and Jenkins, 1970; Hannan, 1970; and Brillinger, 1981).

Black-box continuous transfer function models of the type (4.47) have been used in many cases oriented toward control applications (see, e.g., Webb and Soderstrom, 1985). Ziegler and Nichols (1942) determine parameters in such models from step responses and self-oscillatory modes (see Section 6.1). A survey of continuous-time models and their estimation is given in Young (1981).

State-space models in innovations forms as well as the general forms are treated in standard textbooks on control (e.g., Åström and Wittenmark, 1984). The use of continuous-time representations for estimation using discrete data has been discussed, for example, in Mehra and Tyler (1973) and Åström and Källström (1976). The continuous-time model structure is usually arrived at after an initial modeling step. See, for example, Wellstead (1979) and Nicholson (1981) for general modeling techniques and examples.

Distributed parameter models, and their estimation are treated in, for example, Banks, Crowley, and Kunisch (1983), Kubrusly (1977), Qureshi, Ng, and Goodwin (1980), and Polis and Goodson (1976). Example 4.3 is studied experimentally in Leden, Hamza, and Sheirah (1976).

The prediction aspect of models was emphasized in Ljung (1974, 1978a). Identifiability is discussed in many contexts. A survey is given in Nguyen and Wood (1982). Often identifiability is related to convergence of the parameter estimates. Such definitions are given in Åström and Bohlin (1965), Staley and Yuc (1970), and Tse and Anton (1972). Identifiability in terms of the model structure only was introduced by Bellman and Åström (1970), who called it "structural identifiability." Identifiability definitions in terms of the set $D_T(\mathcal{S}, \mathcal{M})$ [defined by (4.131)] were given in Gustavsson, Ljung, and Söderström (1977). The particular definitions of the concepts of model structure and identifiability given in Section 4.5 are novel.

A more general model structure concept than Definition 4.3 would be to let

D_M be a differentiable manifold (see, e.g., Byrnes, 1976). However, in our treatment that possibility is captured by letting a model set be described as a union of (overlapping) ranges of model structures as in (4.126). This manifold structure for linear systems was first described by Kalman (1974), Hazewinkel and Kalman (1976), and Clark (1976). See also Byrnes and Hurt (1979).

The identifiability analysis of SISO models in Section 4.6 essentially goes back to Åström and Bohlin (1965).

The identifiability of multivariable model structures has been dealt with in numerous articles. See, for example, Kailath (1980), Luenberger (1967), Glover and Willems (1974), Rissanen (1974), Ljung and Rissanen (1976), Guidorzi (1981), Gevers and Wertz (1984), van Overbeek and Ljung (1982), and Correa and Glover (1984).

4.9 PROBLEMS

4G.1. Consider the predictor (4.18). Show that the effect from an erroneous initial condition in $\hat{y}(s|\theta), s \le 0$, is bounded by $c \cdot \mu^t$, where μ is the maximum magnitude of the zeros of $C(z)$.

4G.2. *Colored measurement noise:* Suppose that a state-space representation is given as

$$x(t+1) = A_1(\theta)x(t) + B_1(\theta)u(t) + w_1(t)$$
$$y(t) = C_1(\theta)x(t) + v(t) \tag{4.146}$$

where $\{w_1(t)\}$ is white with variance $\overline{R}_1(\theta)$, but the measurement noise $\{v(t)\}$ is not white. A model for $v(t)$ can, however, be given as

$$v(t) = H(q,\theta)v(t) \tag{4.147}$$

with $\{v(t)\}$ being white noise with variance $R_2(\theta)$ and $H(q,\theta)$ monic. Introduce a state-space representation for (4.147):

$$\xi(t+1) = A_2(\theta)\xi(t) + K(\theta)v(t)$$
$$v(t) = C_2(\theta)\xi(t) + v(t) \tag{4.148}$$

Combine (4.146) and (4.147) into a single representation that complies with the structure (4.81) to (4.82). Determine $R_1(\theta)$, $R_{12}(\theta)$, and $R_2(\theta)$. Note that if $w_1(t)$ is zero then the new representation will be directly in the innovations form (4.88).

4G.3 *Verification of the Steady-State Kalman Filter:* The state-space model (4.81) can be written (suppressing the argument θ and assuming dim $y = 1$)

$$y(t) = G(q)u(t) + v_1(t)$$

where

$$G(q) = C(qI - A)^{-1}B$$
$$v_1(t) = C(qI - A)^{-1}w(t) + v(t)$$

Let $R_{12} = 0$. The spectrum of $\{v_1(t)\}$ then is

$$\Phi_1(\omega) = C(e^{i\omega} \cdot I - A)^{-1}R_1(e^{-i\omega} \cdot I - A^T)^{-1}C^T + R_2$$

using Theorem 2.2. The innovations model (4.88) can be written

$$y(t) = G(q)u(t) + v_2(t)$$

$$v_2(t) = H(q)e(t), \qquad H(q) = C(qI - A)^{-1}K + 1$$

The spectrum of $\{v_2(t)\}$ thus is

$$\Phi_2(\omega) = \lambda[C(e^{i\omega} \cdot I - A)^{-1}K + 1][C(e^{-i\omega} \cdot I - A)^{-1}K + 1]^T$$

where λ is the variance of $e(t)$.

(a) Show by direct calculation that

$$\Phi_1(\omega) - \Phi_2(\omega) \equiv 0$$

utilizing the expressions (4.84) and (4.88b). The two representations thus have the same second-order properties, and if the noises are Gaussian, they are indistinguishable in practice (see Problem 2E.3).

(b) Show by direct calculation that

$$1 - H^{-1}(q) = 1 - [1 + C(qI - A)^{-1}K]^{-1} = C(qI - A + KC)^{-1}K$$

and

$$H^{-1}(q)G(q) = [1 + C(qI - A)^{-1}K]^{-1}C(qI - A)^{-1}B = C(qI - A + KC)^{-1}B$$

(c) Note that the predictor (4.83) can be written as (4.85):

$$\hat{y}(t|\theta) = C(qI - A + KC)^{-1}Bu(t) + C(qI - A + KC)^{-1}Ky(t)$$

and thus that (a) and (b) together with (3.20) constitute a derivation of the steady-state Kalman filter.

4G.4. Consider a model structure $\mathcal{M}$, with predictor function gradient $\Psi(z, \theta)$ defined in (4.118). Define the $d \times d$ matrix

$$\Gamma_1(\theta) = \int_{-\pi}^{\pi} \Psi(e^{i\omega}, \theta)\Psi^T(e^{-i\omega}, \theta)d\omega$$

(a) Show that $\mathcal{M}$ is locally identifiable at θ if $\Gamma_1(\theta)$ is nonsingular.

(b) Let $\mathbf{T}'(z, \theta)$ be defined by (4.122), and let

$$\Gamma_2(\theta) = \int_{-\pi}^{\pi} \mathbf{T}'(e^{i\omega}, \theta)[\mathbf{T}'(e^{-i\omega}, \theta)]^T d\omega$$

Use (4.121) to show that $\Gamma_2(\theta)$ is nonsingular if and only if $\Gamma_1(\theta)$ is. [Note that by assumption $H(q)$ has no zeros on the unit circle.] $\Gamma_2(\theta)$ can thus be used to test local identifiability.

4G.5. Consider an output error structure with several inputs

$$y(t) = \frac{B_1(q)}{F(q)}u_1(t) + \cdots + \frac{B_m(q)}{F(q)}u_m(t) + e(t)$$

Show that this structure is globally identifiable at a value θ^* if and only if there is no common factor to all of the $m + 1$ polynomials

$$z^{n_f}F^*(z), \qquad z^{n_b}B_i^*(z), \qquad i = 1, \ldots, m$$

$$n_f = \text{degree } F^*(z), \qquad n_b = \max \text{degree } B_i^*(z)$$

θ^* here corresponds to the starred polynomials.

4G.6. The Kronecker product of an $m \times n$ matrix $A = (a_{ij})$ and a $p \times r$ matrix $B = (b_{ij})$ is defined as (see, e.g., Barnett, 1975)

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B \ldots a_{1n}B \\ a_{21}B & a_{22}B \ldots a_{2n}B \\ \vdots & \vdots \quad \vdots \\ a_{m1}B & a_{m2}B \ldots a_{mn}B \end{bmatrix}$$

This is an $mp \times nr$ matrix. Define the operator "col" as the operation to form a column vector out of a matrix by stacking its columns on top of each other:

$$\text{col } B = \begin{bmatrix} B^1 \\ B^2 \\ \vdots \\ B^r \end{bmatrix}, \qquad (rp \times 1 \text{ vector})$$

where B^j is the jth column of B.

Consider (4.53) to (4.56). Show that (4.55) can be transformed into (4.56) with

$$\theta = \text{col } \theta^T$$
$$\varphi(t) = \varphi(t) \otimes I_p$$

where I_p is the $p \times p$ unit matrix. Are other variants of θ and φ also possible?

4G.7. Consider the continuous-time state-space model (4.93) to (4.94). Assume that the measurements are made in wideband noise with high variance, idealized as

$$\bar{y}(t) = Hx(t) + \bar{v}(t)$$

where $\bar{v}(t)$ is formal continuous-time white noise with covariance function

$$E\bar{v}(t)\bar{v}^T(s) = \bar{R}_2(\theta)\delta(t - s)$$

Assume that $\bar{v}(t)$ is independent of $\bar{w}(t)$. Let the output be defined as

$$y((k + 1)T) = y_{k+1} = \frac{1}{T}\int_{t=kT}^{(k+1)T} \bar{y}(t)dt$$

Show that the sampled-data system can be represented as (4.95) and (4.96) but with

$$y(kT) = C_T(\theta)x(kT) + D_T(\theta)u(kT-T) + v_T(kT)$$

$$C_T(\theta) = \frac{1}{T}H\Phi_T(\theta)$$

$$Ew_T(kT)v_T^T(kT) = R_{12}(\theta) = \frac{1}{T}\int_0^T e^{F(\theta)\tau}\bar{R}_1(\theta)\Phi_{T-\tau}^T(\theta)H^T d\tau$$

$$Ev_T(kT)v_T^T(kT) = R_2(\theta) = \frac{1}{T}\bar{R}_2(\theta) + \frac{1}{T^2}\int_0^T H\Phi_{T-\tau}(\theta)\bar{R}_1(\theta)\Phi_{T-\tau}^T(\theta)H^T d\tau$$

$$\Phi_T(\theta) = \int_0^T e^{F(\theta)\tau}d\tau; \qquad D_T(\theta) = -\frac{1}{T}\int_0^T H\Phi_\tau(\theta)d\tau$$

A derivation is given in Wahlberg (1985).

4G.8. Consider the ARX model (4.7). Introduce the δ-operator

$$\delta = 1 - q^{-1}$$

and reparametrize the models in terms of coefficients of powers of δ. Work out the details of a second-order example. Such a parametrization has the advantage of being less sensitive to numerical errors when the sampling interval is short (Goodwin, 1985).

4E.1. Consider the ARX model structure

$$y(t) + a_1 y(t-1) + \cdots + a_{n_a} y(t-n_a) = b_1 u(t-1) + \cdots + b_{n_b} u(t-n_b) + e(t)$$

where b_1 is known to be 0.5. Write the corresponding predictor in the linear regression form (4.13).

4E.2. Consider the continuous-time model (4.72) of the dc servo with $T_\ell(t) \equiv 0$. Apply the Euler approximation (2.25) to obtain an approximate discrete-time transfer function that is a simpler function of θ.

4E.3. Consider the small network of tanks in Figure 4.8. Each tank holds 10 volume units of fluid. Through the pipes A and E flow 1 volume unit per second, through the pipe B, α units, and through C and D, 1 − α units per second. The concentration of a certain substance in the fluid is u in pipe A (the input) and y in pipe E (the output). Write down a structured state-space model for this system. Assume that each tank is perfectly mixed (i.e., the substance has the same concentration throughout the tank). (Models of this character are known as *compartmental models* and are very common in chemical and biological applications; see Godfrey, 1983.)

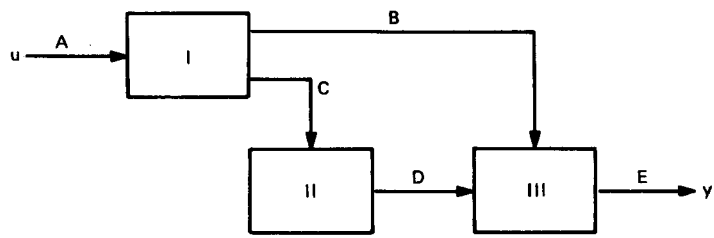

Figure 4.8 A network of tanks.

4E.4. Consider the *RLC* circuit in Figure 4.9 with ideal voltage source $u_v(t)$ and ideal current source $u_i(t)$. View this circuit as a linear time-invariant system with two inputs

$$u(t) = \begin{bmatrix} u_v(t) \\ u_i(t) \end{bmatrix}$$

and one output: the voltage $y(t)$. R, L, and C are unknown constants. Discuss several model set parametrizations that could be feasible for this system and describe their advantages and disadvantages.

Hint: The basic equations for this circuit are

$$u_v(t) = L\frac{di(t)}{dt} + y(t) + R[i(t) + u_i(t)]$$

$$y(t) = \frac{1}{C}\int_0^t i(\tau)d\tau$$

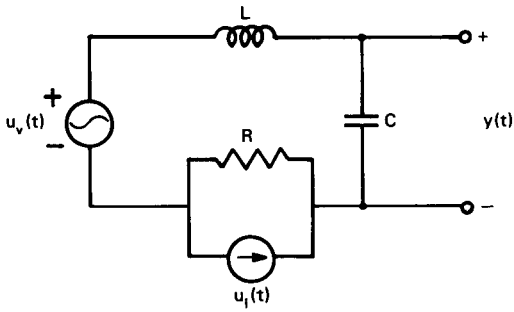

Figure 4.9 A simple circuit.

4E.5. A state-space model of ship-steering dynamics can be given as follows:

$$\frac{d}{dt}\begin{bmatrix} v(t) \\ r(t) \\ h(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} v(t) \\ r(t) \\ h(t) \end{bmatrix} + \begin{bmatrix} b_{11} \\ b_{21} \\ 0 \end{bmatrix} u(t)$$

where $u(t)$ is the rudder angle, $v(t)$ the sway velocity, $r(t)$ the turning rate, and $h(t)$ the heading angle.

(a) Suppose only $u(t)$ and $y(t) = h(t)$ are measured. Show that the six parameters a_{ij}, b_{ij} are not identifiable.

(b) Try also to show that if $u(t)$ and $y(t) = \begin{bmatrix} v(t) \\ h(t) \end{bmatrix}$ are measured then all six parameters are globally identifiable at values such that the model is controllable. If you cannot complete the calculations, indicate how you would approach the problem (reference: Godfrey and DiStefano, 1985).

4E.6. Consider the model structure (4.88) with

$$A(\theta) = \begin{bmatrix} -a_1 & 1 \\ -a_2 & 0 \end{bmatrix}, \qquad B(\theta) = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

$$C(\theta) = [1 \quad 0], \qquad K(\theta) = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$$

$$\theta = [a_1 a_2 b_1 b_2 k_1 k_2]^T, \qquad \theta \in D_1 \subset \mathbf{R}^6$$

and another structure

$$A(\eta) = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \qquad B(\eta) = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

$$C(\eta) = [\gamma_1 \quad \gamma_2], \qquad K(\eta) = \begin{bmatrix} \kappa_1 \\ \kappa_2 \end{bmatrix}$$

$$\eta = [\lambda_1 \lambda_2 \mu_1 \mu_2 \gamma_1 \gamma_2 \kappa_1 \kappa_2]^T, \qquad \eta \in D_2 \subset \mathbf{R}^8$$

Determine D_1 and D_2 so that the two model structures determine the same model set. What about identifiability properties?

4E.7. Consider the heated metal rod of Example 4.3. Introduce a five-state lumped approximation and write down the state-space model explicitly.

4E.8. Consider the OE model structure with $n_b = 2$, $n_f = 1$, and b_1 fixed to unity:

$$y(t) = \frac{q^{-1} + b_2 q^{-2}}{1 + f_1 q^{-1}} u(t) + e(f), \qquad \theta = [b_2 f_1]^T$$

Determine $\Gamma_2(\theta)$ of Problem 4G.4 explicitly. When is it singular?

4E.9. Consider the model structures

$$\mathcal{M}_1: \quad y(t) = -ay(t-1) + bu(t-1)$$
$$\theta = \begin{bmatrix} a \\ b \end{bmatrix}, \qquad D_{\mathcal{M}_1} = \{|a| \le 1, b > 0\}$$

and

$$\mathcal{M}_2: \quad y(t) = -(\cos\alpha)y(t-1) + e^{\beta}u(t-1)$$
$$\eta = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \qquad D_{\mathcal{M}_2} = \{0 \le \alpha \le \pi, -\infty < \beta < \infty\}$$

Show that $\mathcal{R}(\mathcal{M}_1) = \mathcal{R}(\mathcal{M}_2)$. Discuss possible advantages and disadvantages with the two structures.

4E.10. Consider the dc-motor model (4.72). Assume that the torque T_ℓ can be seen as a white-noise zero mean disturbance with variance σ^2 (i.e., the variations in T_ℓ are random and fast compared to the dynamics of the motor). Apply (4.94) to (4.96) to determine $R_1(\theta)$ and $R_{12}(\theta)$ in a sampled model (4.81) and (4.82) of the motor, with $A(\theta)$ and $B(\theta)$ given by (4.74) and

$$\theta = \begin{bmatrix} \tau \\ \beta \\ \gamma \end{bmatrix}, \qquad \gamma = \gamma' \cdot \sigma$$

As an alternative, we could use a directly parametrized innovations form (4.88) with $A(\theta)$ and $B(\theta)$ again given by (4.74), but

$$K(\theta) = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \quad \text{and} \quad \theta = [\tau \ \beta \ k_1 \ k_2]^T$$

Discuss the advantages and disadvantages of these two parametrizations.

4E.11. Consider the system description

$$x(t+1) = ax(t) + bu(t) + \xi(t)$$
$$y(t) = x(t) + e(t)$$

where $e(t)$ is white Gaussian noise and $\xi(t)$ has the distribution

$$\xi(t) = 0, \qquad \text{w.p. } 1 - \lambda$$
$$\xi(t) = +1, \qquad \text{w.p. } \frac{\lambda}{2}$$
$$\xi(t) = -1, \qquad \text{w.p. } \frac{\lambda}{2}$$

The coefficients a, b, and λ are adjustable parameters. Can this description be cast into the form (4.4)? If so, at the expense of what approximations?

4E.12. Consider a multivariable ARX model set

$$y(t) + A_1 y(t-1) + A_2 y(t-2) = B_1 u(t-1) + e(t)$$

where $\dim y = p = 2$, $\dim u = m = 1$, and where the matrices are parametrized as

$$A_1 = \begin{bmatrix} \times & \times \\ \alpha & \times \end{bmatrix}, \quad A_2 = \begin{bmatrix} \times & \times \\ 0 & \times \end{bmatrix}, \quad B_1 = \begin{bmatrix} \beta \\ \times \end{bmatrix}$$

where α and β are known values and $\times$ indicates a parameter to be estimated. Write the predictor in the form

$$\hat{y}(t|\theta) = \varphi^T(t)\theta + \mu(t)$$

with $\mu(t)$ as a known term and give explicit expressions for φ and θ. Can this predictor be written in the form (4.55)?

4T.1. Determine the k-step-ahead predictor for the ARMAX model (4.15).

4T.2. Give an expression for the k-step-ahead predictor for (4.88).

4T.3. Suppose that $W_u(q)$ and $W_y(q)$ are given functions, known to be determined as k-step-ahead predictors for the system description

$$y(t) = G(q)u(t) + H(q)e(t)$$

Can $G(e^{i\omega})$ and $H(e^{i\omega})$ be uniquely computed from $W_u(e^{i\omega})$ and $W_y(e^{i\omega})$? What if G and H are known to be of the ARMAX structure

$$G(q) = \frac{B(q)}{A(q)}, \qquad H(q) = \frac{C(q)}{A(q)}$$

where A, B, and C have known (and suitable) orders?

4D.1. Prove Lemma 4.2.

4S.1. In many of the S-problems aiming at a software package, we shall work with the general I/O-model family (4.34). Select the general family or your favorite subsets of it (see Table 4.1) for these problems and decide on a matrix structure

TH

to contain the polynomials orders *na, nb, nc, nd,* and *nf* and the delay *nk,* as well as the parameter values a_i, b_i, and so on, and the variance λ of the innovations. The variable TH thus uniquely determines a particular model. Reserve space for a covariance matrix of dimension $d \times d$ to be included in TH later. Write a MACRO

PRESENT(TH)

for presenting the model on the screen in user-friendly format. Also write a MACRO

[G,PHIV] = TRF(TH)

that computes the transfer function

$$G(e^{i\omega}) = \frac{B(e^{i\omega})}{A(e^{i\omega})F(e^{i\omega})}, \qquad \omega = \omega_1, \ldots, \omega_N$$

and the noise spectrum

$$\text{PHIV}(\omega) = \lambda \left| \frac{C(e^{i\omega})}{A(e^{i\omega})D(e^{i\omega})} \right|^2, \qquad \omega = \omega_1 \ldots \omega_N$$

and returns them as vectors to be used with the BODEPLOT command.

4S.2. Write a MACRO

$$y = \text{SIMU}(u, e, \text{TH})$$

that simulates (4.34) and a MACRO

$$e = \text{PE}(y, u, \text{TH})$$

that computes the prediction errors (4.37). Here and henceforth, u, e, and y are the vectors u^t, e^t, and y^t. [For future use it is a good idea to let PE also return the sequences (vectors) w and v in (4.38).]

APPENDIX 4A: IDENTIFIABILITY OF BLACK-BOX MULTIVARIABLE MODEL STRUCTURES

The topic of multivariable model structures and canonical forms for multivariable systems is often regarded as difficult, and there is an extensive literature in the field. We shall here give a no-frills account of the problem, and the reader is referred to the literature for more insights and deeper results. See the bibliography.

The issue still is whether (4.133) holds at a given θ. Our development parallels the one in Section 4.6. We start by discussing polynomial parametrizations or MFDs, such as (4.52) to (4.58), and then turn to state-space models. Throughout the section, p denotes the number of outputs and m the number of inputs.

Matrix Fraction Descriptions (MFD)

Consider first the simple multivariable ARX structure (4.49) or (4.53). This uses

$$G(z,\theta) = A^{-1}(z)B(z), \qquad H(z,\theta) = A^{-1}(z) \tag{4A.1}$$

with θ comprising all the coefficients of the matrix polynomials (in $1/z$) $A(z)$ and $B(z)$. These could be of arbitrary orders. Just as for the SISO case (4.136), it is immediate to verify that (4.133) holds for all θ^*. Hence the model structure given by the MFD (4A.1) is *strictly globally identifiable.*

Let us now turn to the output error model structure

$$G(z,\theta) = F^{-1}(z)B(z), \qquad H(z,\theta) = I \tag{4A.2}$$

It should be noted that the analysis of (4A.2) contains also the analysis of the multivariable ARMAX structure and multivariable Box–Jenkins models. See the corollary to Theorem 4A.1, which follows.

The matrix polynomial $F(z)$ is here a $p \times p$ matrix

$$F(z) = \begin{bmatrix} F_{11}(z) & F_{12}(z) \dots F_{1p}(z) \\ F_{21}(z) & F_{22}(z) \dots F_{2p}(z) \\ \vdots & \vdots \quad \vdots \\ F_{p1}(z) & F_{p2}(z) \dots F_{pp}(z) \end{bmatrix} = F^{(0)} + F^{(1)} z^{-1} + \cdots + F^{(\nu)} z^{-\nu} \tag{4A.3}$$

whose entries are polynomials in z^{-1}:

$$F_{ij}(z) = f_{ij}^{(0)} + f_{ij}^{(1)} z^{-1} + \cdots + f_{ij}^{(\nu_{ij})} z^{-\nu_{ij}} \tag{4A.4}$$

The degree of the F_{ij} polynomial will thus be denoted by ν_{ij} and $\nu = \max \nu_{ij}$. Similarly, $B(z)$ is a $p \times m$ matrix polynomial. Let the degrees of its entries be denoted by μ_{ij}.

The structure issue is really to select the orders ν_{ij} and μ_{ij} [i.e., $p(p+m)$ integers]. This will give a staggering amount of possible model structures. Some special cases discussed in the literature are

1. $\nu_{ij} = n$, $\mu_{ij} = r$ (4A.5)
2. $\nu_{ij} = 0, i \neq j$; $\quad \nu_{ii} = n_i$, $\mu_{ij} = r_i$ (4A.6)
3. $\nu_{ij} = n_j$, all i; $\quad \mu_{ij} = r_j$, all i (4A.7)

In all these cases we fix the leading matrix to be a unit matrix:

$$F^{(0)} = I; \qquad \text{i.e.,} f_{ij}^{(0)} = \delta_{ij} \tag{4A.8}$$

The form (4A.5) is called the "full polynomial form" in Söderström and Stoica (1983). It clearly is a special case of (4A.7). It is used and discussed in Hannan (1969, 1976), Kashyap and Rao (1976), Jakeman and Young (1979), and elsewhere.

The form (4A.6) gives a diagonal F-matrix and has been used, for example, in Kashyap and Nasburg (1974), Sinha and Caines (1977), and Gauthier and Landau (1978).

The structure (4A.7) where the different columns are given different orders is discussed, for example, in Guidorzi (1975), Gauthier and Landau (1978), and Gevers and Wertz (1984).

Remark. In the literature, especially the one discussing canonical forms rather than identification applications, often the polynomials

$$\overline{F}(z) = z^{\nu} F(z) = F^{(0)} z^{\nu} + F^{(1)} z^{\nu - 1} + \cdots + F^{(\nu)} \tag{4A.9}$$

in the variable z are considered instead of $F(z)$ (just as we did the SISO case). Canonical representations of $\overline{F}(z)$ [such as the "Hermite form"; see Dickinson, Kailath, and Morf, 1974; Hannan, 1971a; or Kailath, 1980] will then typically involve singular matrices $F^{(0)}$. Such representations are not suitable for our purposes since $y(t)$ cannot be solved for explicitly in terms of past data. ■

The identifiability properties of the diagonal form (4A.6) can be analyzed by SISO arguments. For the others we need some theory for matrix polynomials.

Some Terminology for Matrix Polynomials

Kailath (1980), Chapter 6, gives a detailed account of various concepts and properties of matrix polynomials. We shall here need just a few:

A $p \times p$ matrix polynomial $P(x)$ is said to be *unimodular* if $\det P(x) =$ constant. Then $P^{-1}(x)$ is also a matrix polynomial. Two polynomials $P(x)$ and $Q(x)$ with the same number of rows have a *common left divisor* if there exists a matrix polynomial $L(x)$ such that

$$P(x) = L(x)\tilde{P}(x)$$

$$Q(x) = L(x)\tilde{Q}(x)$$

for some matrix polynomials $\tilde{P}(x)$ and $\tilde{Q}(x)$.

$P(x)$ and $Q(x)$ are said to be *left coprime* if all common left divisors are unimodular. This is a direct extension of the corresponding concept for scalar polynomials. A basic theorem says that if $P(x)$ and $Q(x)$ are left coprime then there exist matrix polynomials $A(x)$ and $B(x)$ such that

$$P(x)A(x) + Q(x)B(x) = I \quad \text{(identity matrix)} \tag{4A.10}$$

Loss of Identifiability in Multivariable MFD Structures

We can now state the basic identifiability result.

Theorem 4A.1. Consider the output error MFD model structure (4A.2) with the polynomial degrees chosen according to the scheme (4A.7). Let θ comprise all the coefficients in the resulting matrix polynomials, and let $F_*(z)$ and $B_*(z)$ be the polynomials in $1/z$ that correspond to the value θ^*. Let

$$D_p(z) = \operatorname{diag}(z^{n_1}, \ldots, z^{n_p})$$

$$D_m(z) = \operatorname{diag}(z^{r_1}, \ldots, z^{r_m})$$

be diagonal matrices, with n_i and r_i defined in (4A.7), and define $\tilde{F}_*(z) = F_*(z)D_p(z)$, $\tilde{B}_*(z) = B_*(z)D_m(z)$ as polynomials in z. Then the model structure in question is globally and locally identifiable at θ^* if and only if

$$\tilde{F}_*(z) \text{ and } \tilde{B}_*(z) \text{ are left coprime} \tag{4A.11}$$

Proof. Let θ correspond to $F(z)$ and $B(z)$, and assume that

$$G(z, \theta) = G(z, \theta^*) = F^{-1}(z)B(z) = F_*^{-1}(z)B_*(z)$$

This can also be written as

$$D_p(z)\tilde{F}^{-1}(z)\tilde{B}(z)D_m^{-1}(z) = D_p(z)\tilde{F}_*^{-1}(z)\tilde{B}_*(z)D_m^{-1}(z)$$

where $\tilde{F}$ and $\tilde{B}$ are defined analogously to $\tilde{F}_*$ and $\tilde{B}_*$. This gives

$$\tilde{B}_*(z) = \tilde{F}_*(z)\tilde{F}^{-1}(z)\tilde{B}(z) \tag{4A.12}$$

When $\tilde{B}_*$ and $\tilde{F}_*$ are left coprime there exist, according to (4A.10), matrix polynomials $X(z)$ and $Y(z)$ such that

$$\tilde{F}_*(z)X(z) + \tilde{B}_*(z)Y(z) = I$$

Inserting (4A.12) gives

$$\tilde{F}_*(z)\tilde{F}^{-1}(z)[\tilde{F}(z)X(z) + \tilde{B}(z)Y(z)] = I$$

or

$$\tilde{F}(z)X(z) + \tilde{B}(z)Y(z) = \tilde{F}(z)\tilde{F}_*^{-1}(z) \triangleq U(z)$$

Since the left side is a matrix polynomial in z, so is $U(z)$. We have

$$\tilde{F}(z) = U(z)\tilde{F}_*(z) \tag{4A.13}$$

Note that, by (4A.8),

$$I = \lim_{z\to\infty} F(z) = \lim_{z\to\infty} \tilde{F}(z)D_p^{-1}(z) = \lim_{z\to\infty} \tilde{F}_*(z)D_p^{-1}(z)$$

Hence, multiplying (4A.13) by $D_p^{-1}(z)$ gives

$$I = \lim_{z\to\infty} U(z)$$

which since $U(z)$ is a polynomial in z, shows that $U(z) \equiv I$, and hence $F(z) \equiv F_*(z)$, which in turn implies that $B(z) = B_*(z)$, and the if-part of the theorem has been proved. If (4A.11) does not hold, a common, nonunimodular, left factor $U_*(z)$ can be pulled out from $F_*(z)$ and $B_*(z)$ and be replaced by an arbitrary matrix with the same orders as $U_*(z)$ [subject to the constraint (4A.8)]. This proves the only-if-part of the theorem. ■

The theorem can immediately be extended to a model structure

$$G(z,\theta) = F^{-1}(z)B(z), \qquad H(z,\theta) = D^{-1}(z)C(z) \tag{4A.14}$$

with F and D subject to the degree structure (4A.7). It can also be extended to the multivariable ARMAX structure:

$$G(z,\theta) = A^{-1}(z)B(z), \qquad H(z,\theta) = A^{-1}(z)C(z) \tag{4A.15}$$

Corollary 4A.1. Consider the ARMAX model structure (4A.15) with the degrees of the polynomial $A(z)$ subject to (4A.7). Let $\tilde{A}_*(z)$ and $\tilde{\beta}_*(z) = [\tilde{B}_*(z)\ \tilde{C}_*(z)]$, a $p \times (m+p)$ matrix polynomial, be the polynomials that correspond to θ^*, as described in the theorem. Then the structure is identifiable at θ^* if and only if

$$\tilde{A}_*(z) \text{ and } \tilde{\beta}_*(z) \text{ are left coprime} \quad ■$$

The usefulness of these identifiability results lies in the fact that only p orders (the column degrees) have to be chosen with care to find a suitable identifiable

structure, despite the fact that $p \cdot m$ [or even $p \cdot (m + p)$ in the ARMAX case] different transfer functions are involved.

State-space Model Structures

For a multivariable state-space model (4.143a), we introduce a parametric structure, analogous to (4.143b):

$$A(\theta) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \times & \times & \times & \times & \times & \times & \times & \times & \times \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \times & \times & \times & \times & \times & \times & \times & \times & \times \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \times & \times & \times & \times & \times & \times & \times & \times & \times \end{bmatrix}, \quad B(\theta) = \begin{bmatrix} \times & \times \\ \times & \times \\ \times & \times \\ \times & \times \\ \times & \times \\ \times & \times \\ \times & \times \\ \times & \times \\ \times & \times \end{bmatrix}$$

$$K(\theta) = \begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix}, \quad C(\theta) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{4A.16}$$

The number of rows with $\times$'s in $A(\theta)$ equals the number of outputs. We have thus illustrated the structure for $n = 9$, $p = 3$, $m = 2$. In words, the general structure can be defined as:

> Let $A(\theta)$ initially be a matrix filled with zeros and with ones along the superdiagonal. Let then row numbers $r_1, r_2, \ldots, r_p$, where $r_p = n$, be filled with parameters. Take $r_0 = 0$ and let $C(\theta)$ be filled with zeros, and then let row i have a one in column $r_{i-1} + 1$. Let $B(\theta)$ and $K(\theta)$ be filled with parameters. (4A.17)

The parametrization is uniquely characterized by the p numbers r_i that are to be chosen by the user. We shall also use

$$\nu_i = r_i - r_{i-1}$$

and call

$$\bar{\nu}_n = \{\nu_1, \ldots, \nu_p\} \tag{4A.18}$$

the *multiindex* associated with (4A.17). Clearly,

$$n = \sum_{i=1}^{p} \nu_i \tag{4A.19}$$

By a *multiindex* $\bar{\nu}_n$ we henceforth understand a collection of p numbers $\nu_i \geq 1$ subject to (4A.19). For given n and p, there exist $\binom{n}{p}$ different multi-indexes. Notice that the structure (4A.17) contains $2np + mn$ parameters regardless of $\bar{\nu}_n$.

The key property of a "canonical" parametrization like (4A.16) is that the corresponding state vector $x(t, \theta)$ can be interpreted in a pure input–output context. This can be seen as follows. Fix time t, and assume that $u(s) = e(s) \equiv 0$ for $s \geq t$. Denote the corresponding outputs that are generated by the model for $s \geq t$ by $\hat{y}_\theta(s|t-1)$. We could think of them as projected outputs for future times s as calculated at time $t-1$. The state-space equations give directly

$$\begin{aligned}
\hat{y}_\theta(t|t-1) &= C(\theta)x(t,\theta) \\
\hat{y}_\theta(t+1|t-1) &= C(\theta)A(\theta)x(t,\theta) \\
&\vdots \\
\hat{y}_\theta(t+n-1|t-1) &= C(\theta)A^{n-1}(\theta)x(t,\theta)
\end{aligned} \tag{4A.20}$$

With

$$\mathbb{O}_n(\theta) = \begin{bmatrix} C(\theta) \\ C(\theta)A(\theta) \\ \vdots \\ C(\theta)A^{n-1}(\theta) \end{bmatrix} \tag{4A.21}$$

(the $np \times n$ observability matrix) and

$$\hat{Y}_n^\theta(t) = \begin{bmatrix} \hat{y}_\theta(t|t-1) \\ \vdots \\ \hat{y}_\theta(t+n-1|t-1) \end{bmatrix}$$

We can write (4A.20) as

$$\hat{Y}_n^\theta(t) = \mathbb{O}_n(\theta)x(t,\theta) \tag{4A.22}$$

It is straightforward to verify that (4A.17) has a fundamental property: The $np \times n$ observability matrix $\mathbb{O}_n(\theta)$ will have n rows that together constitute the unit matrix, regardless of θ. The reader is invited to verify that row number $kp + i$ of $\mathbb{O}_n$ will be

$$[00\ldots010\ldots0]$$

with 1 in position $r_{i-1} + k + 1$. This holds for $1 \leq i \leq p$, $0 \leq k < \nu_i$. Thus (4A.22) implies that the state variables corresponding to the structure (4A.17) are

$$x_{r_{i-1}+k+1}(t,\theta) = \hat{y}_\theta^{(i)}(t+k|t-1), \qquad i = 1,\ldots,p, \qquad 0 \leq k < \nu_i \tag{4A.23}$$

Here superscript (i) denotes the ith component of y. This interpretation of state

variables as predictors is discussed in detail in Akaike (1974b) and Rissanen (1974). By the relation (4A.23), n rows are picked out from the np vector $\hat{Y}_n^{\theta}(t)$ in (4A.22). The indexes of these rows are uniquely determined by the multiindex $\bar{\nu}_n$. Let them be denoted by

$$I_{\bar{\nu}_n} = \{(k-1)p + i; \quad 1 \le k \le \nu_i; \quad 1 \le i \le p\} \tag{4A.24}$$

The key relationship is (4A.23). It shows that the state variables depend only on the input–output properties of the associated model.

Consider now two values θ^* and θ that give the same input–output properties of (4A.17). Then $\hat{y}_{\theta}(t+k|t-1) = \hat{y}_{\theta^*}(t+k|t-1)$, since these are computed from input–output properties only. Thus $x(t,\theta) = x(t,\theta^*)$. Now, if θ^* corresponds to a minimal realization, so must θ, and Theorem 6.2.4 of Kailath (1980) gives that there exists an invertible matrix T such that

$$\begin{aligned} A(\theta^*) &= TA(\theta)T^{-1}, \qquad B(\theta^*) = TB(\theta) \\ K(\theta^*) &= TK(\theta), \qquad C(\theta^*) = C(\theta)T^{-1} \end{aligned} \tag{4A.25}$$

corresponding to the change of basis

$$x(t,\theta^*) = Tx(t,\theta) \tag{4A.26}$$

But (4A.26) together with our earlier observation that $x(t,\theta^*) = x(t,\theta)$ shows that $T = I_n$ and hence that $\theta^* = \theta$.

We have now proved the major part of the following theorem:

Theorem 4A.2. Consider the state-space model structure (4A.17). This structure is globally and locally identifiable at θ^* if and only if $\{A(\theta^*), [B(\theta^*)\ K(\theta^*)]\}$ is controllable.

Proof. The if-part was proved previously. To show the only-if-part, we find that if θ^* does not give a controllable system then its input–output properties can be described by a lower-dimensional model with an additional, arbitrary, non-controllable model. This can be accomplished by infinitely many different θ's. ■

It follows from the theorem that the parametrization (4A.17) is globally identifiable, and as such is a good candidate to describe systems of order n. What is not clear yet is whether *any* nth-order linear system can be represented in the form (4A.17) for *an arbitrary choice* of multiindex $\bar{\nu}_n$. That is the question we now turn to.

Hankel-Matrix Interpretation

Consider a multivariable system description

$$y(t) = G_0(q)u(t) + H_0(q)e(t) = T_0(q)\chi(t) \tag{4A.27}$$

with

$$T_0(q) = [G_0(q)\ H_0(q)], \qquad \chi(t) = \begin{bmatrix} u(t) \\ e(t) \end{bmatrix}$$

Assume that $T_0(q)$ has full row rank [i.e., $LT_0(q)$ is not identically zero for any nonzero $1 \times p$ vector L]. Let

$$T_0(q) = [0 \quad I] + \sum_{k=1}^{\infty} H_k q^{-k} \tag{4A.28}$$

be the impulse response of the system. The matrices H_k are here $p \times (p + m)$. Define the matrix

$$\mathcal{H}_{r,s} = \begin{bmatrix} H_1 & H_2 & \dots & H_s \\ H_2 & H_3 & \dots & H_{s+1} \\ H_3 & H_4 & \dots & H_{s+2} \\ \vdots & \vdots & & \vdots \\ H_r & H_{r+1} & \dots & H_{r+s-1} \end{bmatrix} \tag{4A.29}$$

This structure with the same block elements along antidiagonals is known as a *block Hankel matrix*. Consider the semifinite matrix $\mathcal{H}_r = \mathcal{H}_{r,\infty}$. For this matrix we have the following two fundamental results.

Lemma 4A.1. Suppose that the n rows $I_{\bar{\nu}_n}$ [see (4A.24)] of $\mathcal{H}_n$ span all the rows of $\mathcal{H}_{n+1}$. Then the system (4A.27) can be represented in the state-space form (4A.17) corresponding to the multiindex $\bar{\nu}_n$. ■

The proof consists of an explicit construction and is given at the end of this appendix.

Lemma 4A.2. Suppose that

$$\operatorname{rank} \mathcal{H}_{n+1} \leq n \tag{4A.30}$$

Then there exists a multiindex $\bar{\nu}_n$ such that the n rows $I_{\bar{\nu}_n}$ span $\mathcal{H}_{n+1}$. The proof of this lemma is also given at the end of this appendix. ■

It follows from the two lemmas that (4A.30) is a sufficient condition for (4A.27) to be an n-dimensional linear system (i.e., to admit a state-space representation of order n). It is, however, well known that this is also a necessary condition. ($\mathcal{H}$ is obtained as the product of the observability and controllability matrices.) We thus conclude:

Any linear system that can be represented in state-space form of order n can also be represented in the particular form (4A.17) for some multiindex $\bar{\nu}_n$. (4A.31)

When (4A.30) holds, we thus find that the np rows of $\mathcal{H}_n$ span an n-

dimensional (or less) linear space. The *generic* situation is then that the same space is spanned by *any* subset of n rows of $\mathcal{H}_n$. (By this term we mean if we randomly pick from a uniform distribution np row vectors to span an n-dimensional space the probability is 1 that *any* subset of n vectors will span the same space.) We thus conclude:

A state-space representation in the form (4A.17) for a particular multiindex $\bar{\nu}_n$ is capable of describing *almost all* n-dimensional linear systems. (4A.32)

Overlapping Parametrizations

Let $\mathcal{M}_{\bar{\nu}_n}$ denote the model structure (4A.17) corresponding to $\bar{\nu}_n$. The result (4A.31) then implies that the model set

$$\overline{\mathcal{M}} = \bigcup_{\bar{\nu}_n} \mathcal{R}(\mathcal{M}_{\bar{\nu}_n}) \tag{4A.33}$$

(union over all possible multiindices $\bar{\nu}_n$) covers all linear n-dimensional systems. We have thus been able to describe the set of all linear n-dimensional systems as the union of ranges of identifiable structures [cf. (4.126)]. From (4A.32), it follows that the ranges of $\mathcal{M}_{\bar{\nu}_n}$ overlap considerably. This is no disadvantage for identification; on the contrary, one may then change from one structure to another without losing information. The practical use of such overlapping parametrizations for identification is discussed in van Overbeek and Ljung (1982). Using a topological argument, Delchamps and Byrnes (1982) give estimates on the number of overlapping structures needed in (4A.33). See also Hannan and Kavalieris (1984).

Connections Between Matrix Fraction and State-Space Descriptions

In the SISO case the connection between a state-space model in observability form and the corresponding ARMAX model is simple and explicit (see Example 4.2). Unfortunately, the situation is much more complex in the multivariable case. We refer to Gevers and Wertz (1984), Guidorzi (1981), and Beghelli and Guidorzi (1983) for detailed discussions.

We may note, though, the close connection between the indexes ν_i used in (4A.17) and the column degrees n_i in (4A.7). Both determine the number of time shifts of the ith component of y that are explicitly present in the representations. The shifts are, however, forward for the state space and backward for the MFD. The relationship between the ν_i and the observability indexes is sorted out in the proof of Lemma 4A.2.

A practical difference between the two representations is that the state-space representation naturally employs the state $x(t)$ (n variables) as a memory vector for simulation and other purposes. When (4A.2) is simulated in a straightforward fashion, the different delayed components of y and u are stored, a total number of $np + m \cdot \sum r_i$ variables. This is of course not necessary, but an efficient organization

of the variables to be stored amounts to a state-space representation. There are consequently several advantages associated with state-space representations for multivariable systems.

Proofs of Lemmas 4A.1 and 4A.2

It now remains only to prove Lemmas 4A.1 and 4A.2.

Proof of Lemma 4A.1. Let

$$S(t) = \begin{bmatrix} \chi(t-1) \\ \chi(t-2) \\ \vdots \end{bmatrix}$$

Let [cf. (4A.20) to (4A.22)]

$$\hat{y}_0(t+k|t-1) = \sum_{\ell=k+1}^{\infty} H_\ell \chi(t-\ell) \tag{4A.34}$$

and

$$\hat{Y}_N(t) = \begin{bmatrix} \hat{y}_0(t|t-1) \\ \vdots \\ \hat{y}_0(t+N-1|t-1) \end{bmatrix}$$

Then, from (4A.28) and (4A.29),

$$\hat{Y}_N(t) = \mathcal{H}_N S(t) \tag{4A.35}$$

Now enumerate the row indexes i_r of $I_{\bar{\nu}_n}$ in (4A.24) as follows:

$$\begin{aligned} i_1 &= 1, & i_2 &= p+1, \ldots, i_{\nu_1} = (\nu_1 - 1)\cdot p + 1 \\ i_{\nu_1+1} &= 2, & i_{\nu_1+2} &= p+2, \ldots, i_{r_2} = (\nu_2 - 1)\cdot p + 2 \\ &\vdots & & \\ &\text{etc.} & & \\ &\vdots & & \\ i_{r_{p-1}+1} &= p, & i_{r_{p-1}+2} &= p+p, \ldots, i_{r_p} = (\nu_p - 1)\cdot p + p \end{aligned} \tag{4A.36}$$

Recall that

$$r_k = \sum_1^k \nu_j$$

Now construct the n-vector $x(t)$ by taking its rth component to be the i_rth component of $\hat{Y}_N(t)$. Let us now focus on the components $i_1 + p, i_2 + p, \ldots, i_n + p$ of

(4A.35). Collect these components into a vector $\xi(t+1)$. They all correspond to rows of $\mathcal{H}_{n+1}$. But this matrix is spanned by $x(t)$ by the assumption of the lemma. Hence

$$\xi(t+1) = Fx(t) \tag{4A.37}$$

for some matrix F. Now several of the components of $\xi(t+1)$ will also belong to $x(t)$, as shown in (4A.36). The corresponding rows of F will then be zeros everywhere except for a 1 in one position. A moment's reflection on (4A.36) shows that the matrix F will in fact have the structure (4A.17). Also, with H given by (4A.17),

$$y(t) = Hx(t) + e(t) \tag{4A.38}$$

Let us now return to (4A.37). Consider component r of $x(t+1)$, which by definition equals row i_r of $\hat{Y}_N(t+1)$. This row is given as $\hat{y}_0^{(j)}(t+k|t)$ for some values j and k that depend on i_r. But, according to (4A.34), we have

$$\hat{y}_0(t+k|t) = \hat{y}_0(t+k|t-1) + H_k\,\chi(t) \tag{4A.39}$$

Hence

$$x_r(t+1) = \hat{y}_0^{(j)}(t+k|t) = \hat{y}_0^{(j)}(t-1+(k+1)|t-1) + [H_k\,\chi(t)]_j$$

But the first term of the right side equals component number $i_r + p$ of $\hat{Y}_N(t)$ [i.e., $\xi_r(t+1)$]. Hence

$$x(t+1) = \xi(t+1) + M\chi(t) \tag{4A.40}$$

for some matrix M. Equations (4A.37), (4A.38), and (4A.40) now form a state-space representation of (4A.27) within the structure (4A.17) and the lemma is proved. ■

Proof of Lemma 4A.2. The defining property of the Hankel matrix $\mathcal{H}_N$ in (4A.29) means that the same matrix is obtained by either deleting the first block column (and the last block row) or by deleting the first block row. This implies that, if row i of block row k [i.e., row $(k-1)p+i$] lies in the linear span of all rows above it, then so must row i of block $k+1$.

Now suppose that

$$\operatorname{rank} \mathcal{H}_{n+1} = n$$

and let us search the rows from above for a set of linearly independent ones. A row that is not linearly dependent on the ones above it is thus included in the basis; the others are rejected. When the search is finished, we have selected n rows from $\mathcal{H}_{n+1}$. The observation mentioned previously implies that, if row $kp+i$ is included in this basis for $k \geq 1$, then so is row $(k-1)p+i$. Hence the row indexes will obey the structure

$$\begin{matrix} 1, & p+1, & 2p+1, \ldots, (\sigma_1 - 1)p + 1 \\ 2, & p+2, & 2p+2, \ldots, (\sigma_2 - 1)p + 2 \\ p, & p+p, & 2p+p, \ldots, (\sigma_p - 1)p + p \end{matrix}$$

for some numbers $\{\sigma_i\}$ that are known as the *observability indexes* of the system. Since the total number of selected rows is n, we have

$$\sum_{1}^{p} \sigma_i = n$$

The rows thus correspond to the multiindex $\overline{\sigma}_n$ as in (4A.24) and the lemma is proved. Notice that several other multiindexes may give a spanning set of rows; one does not have to look for the first linearly independent rows. ■

5

MODELS FOR TIME-VARYING AND NONLINEAR SYSTEMS

While linear, time-invariant models no doubt form the most common way of describing a dynamical system, it is also quite often useful or necessary to employ other descriptions. In this chapter we shall discuss linear, time-varying models as well as various nonlinear models. We shall also give a formal account of what we mean by a model in general, thus complementing the discussion in Section 4.5 on general linear models.

5.1 LINEAR TIME-VARYING MODELS

Weighting Function

In Chapter 2 we defined a linear system as one where a linear combination of inputs leads to the same linear combination of the corresponding outputs. A general linear system can then be described by

$$y(t) = \sum_{k=1}^{\infty} g_t(k)u(t-k) + v(t) \tag{5.1}$$

If we write

$$g_t(k) = \bar{g}(t, t-k)$$

we find that

$$y(t) = \sum_{s=-\infty}^{t-1} \tilde{g}(t, s)u(s) + v(t) \tag{5.2}$$

where $\tilde{g}(t, s)$, $t = s, s + 1, \ldots$, is the response at time t to a unit input pulse at time s. The function $\tilde{g}(t, s)$ is also known as the *weighting function,* since it describes the weight that the input at time s has in the output at time t.

The description (5.1) is quite analogous to the time-invariant model (2.8), except that the sequence $g_t(k)$ carries the time index t. In general, we could introduce a time-varying transfer function by

$$G_t(q) = \sum_{k=1}^{\infty} g_t(k)q^{-k} \tag{5.3}$$

and repeat most of the discussion in Section 4.2 for time-varying transfer functions. In practice, though, it is easier to deal with time variation in state-space forms.

Time-Varying State-Space Model

Time variation in state-space models (4.88) is simply obtained by letting the matrices be time varying:

$$\begin{aligned} x(t + 1, \theta) &= A_t(\theta)x(t, \theta) + B_t(\theta)u(t) + K_t(\theta)e(t) \\ y(t) &= C_t(\theta)x(t, \theta) + e(t) \end{aligned} \tag{5.4}$$

The predictor corresponding to (4.83) then becomes

$$\begin{aligned} \hat{x}(t + 1, \theta) &= [A_t(\theta) - K_t(\theta)C_t(\theta)]\hat{x}(t, \theta) + B_t(\theta)u(t) + K_t(\theta)y(t) \\ \hat{y}(t|\theta) &= C_t(\theta)\hat{x}(t, \theta) \end{aligned} \tag{5.5}$$

Notice that this can be written

$$\hat{y}(t|\theta) = \sum_{k=1}^{\infty} w_t^u(k, \theta)u(t - k) + \sum_{k=1}^{\infty} w_t^y(k, \theta)y(t - k) \tag{5.6}$$

with

$$\begin{aligned} w_t^u(k, \theta) &= C_t(\theta) \prod_{j=t-k}^{t-1} [A_j(\theta) - K_j(\theta)C_j(\theta)]B_{t-k}(\theta) \\ w_t^y(k, \theta) &= C_t(\theta) \prod_{j=t-k}^{t-1} [A_j(\theta) - K_j(\theta)C_j(\theta)]K_{t-k}(\theta) \end{aligned} \tag{5.7}$$

Similarly, we could start with a time-varying model like (4.81) and (4.82), where the matrices A, B, C, R_1, R_{12}, and R_2 are functions of t. The corresponding predictor will then be given by (4.91) and (4.92).

Two common problems associated with time-invariant systems in fact lead to time-varying descriptions: *nonequal sampling intervals* and *linearization.* If the

system (4.59) and (4.60) is sampled at time instants $t = t_k, k = 1, 2, \ldots$, we can still apply the sampling formulas (4.63) to (4.65) to go from t_k to t_{k+1}, using $T_k = t_{k+1} - t_k$. If this sampling interval is not constant, (4.64) will be a time-varying system. A related case is when *different variables are sampled at different rates.* Then the $C_t(\theta)$ matrix in (5.4) will be time varying in order to pick out the states that are sampled at instant t.

Linearization of Nonlinear Systems

Perhaps the most common use of time-varying linear systems is related to linearization of a nonlinear system around a certain trajectory. Suppose that a nonlinear system is described by

$$\begin{aligned} x(t+1) &= f(x(t), u(t)) + r(x(t), u(t)) \cdot w(t) \\ y(t) &= h(x(t)) + m(x(t), u(t)) \cdot v(t) \end{aligned} \tag{5.8}$$

Suppose also that the disturbance terms $\{w(t)\}$ and $\{v(t)\}$ are white and small, and that the nominal, disturbance-free ($w(t) \equiv 0$; $v(t) \equiv 0$) behavior of the system corresponds to an input sequence $u^*(t)$ and corresponding trajectory $x^*(t)$. Neglecting nonlinear terms, the differences

$$\begin{aligned} \Delta x(t) &= x(t) - x^*(t) \\ \Delta y(t) &= y(t) - h(x^*(t)) \\ \Delta u(t) &= u(t) - u^*(t) \end{aligned}$$

are then subject to

$$\begin{aligned} \Delta x(t+1) &= F(t)\Delta x(t) + G(t)\Delta u(t) + \bar{w}(t) \\ \Delta y(t) &= H(t)\Delta x(t) + \bar{v}(t) \end{aligned} \tag{5.9}$$

where

$$F(t) = \frac{\partial}{\partial x} f(x, u)\Big|_{x^*(t), u^*(t)}, \qquad G(t) = \frac{\partial}{\partial u} f(x, u)\Big|_{x^*(t), u^*(t)}$$

$$H(t) = \frac{\partial}{\partial x} h(x)\Big|_{x^*(t)}$$

Here we have neglected cross terms with the disturbance term (like $\Delta x \cdot v$), in view of our assumption of small disturbances. In (5.9), $\bar{w}(t)$ and $\bar{v}(t)$ are white disturbances with the following covariance properties:

$$\begin{aligned} R_1(t) &= E\bar{w}(t)\bar{w}^T(t) = r(x^*(t), u^*(t))Ew(t)w^T(t)r^T(x^*(t), u^*(t)) \\ R_2(t) &= E\bar{v}(t)\bar{v}^T(t) = m(x^*(t), u^*(t))Ev(t)v^T(t)m^T(x^*(t), u^*(t)) \\ R_{12}(t) &= r(x^*(t), u^*(t))Ew(t)v^T(t)m^T(x^*(t), u^*(t)) \end{aligned} \tag{5.10}$$

This model is now a linear, time-varying, approximate description of (5.8) in a vicinity of the nominal trajectory.

5.2 NONLINEAR MODELS AS LINEAR REGRESSIONS

A nonlinear relationship between the input sequence and the output sequence as in (5.8) clearly gives much richer possibilities to describe systems. At the same time, the situation is far too flexible to allow for definite inference from finite data records: Even a first-order ($\dim x = 1$) model (5.8) without disturbances is specified only up to members in a general infinite-dimensional function space [functions $f(\cdot,\cdot)$ and $h(\cdot)$], while the corresponding linear model is characterized in terms of two real numbers. Various possibilities of parametrizing the functions f and h in general terms (e.g., polynomial approximations) exist, but in most cases it is necessary to have some insight into the character of the nonlinearities to be able to create reasonable model structures. In this section we shall describe how such simple structures can be constructed, provided sufficient physical knowledge of the process to be identified is at hand.

A Linear Regression Structure

In (4.12) we defined a linear regression as a model structure where the prediction is linear in the parameters:

$$\hat{y}(t|\theta) = \varphi^T(t)\theta \tag{5.11}$$

To describe a linear difference equation, the components of the vector $\varphi(t)$ (i.e., the regressors) were chosen as lagged input and output values; see (4.11). When using (5.11) it is, however, immaterial how $\varphi(t)$ is formed; what matters is that it is a known quantity at time t. We can thus let it contain arbitrary transformations of measured data. Let, as usual, y^t and u^t denote the input and output sequences from $s = 1$ to $s = t$. Then we could write

$$\hat{y}(t|\theta) = \theta_1 \varphi_1(u^t, y^{t-1}) + \cdots + \theta_d \varphi_d(u^t, y^{t-1}) = \varphi^T(t)\theta \tag{5.12}$$

with arbitrary functions φ_i of past data. The structure (5.12) could be regarded as a finite-dimensional parametrization of a general, unknown nonlinear predictor. The key is how to choose the functions $\varphi_i(u^t, y^{t-1})$, and this is where physical insight into the system is required. The issue is best illustrated by an example.

Example 5.1 A Solar-heated House

Consider the problem to identify the dynamics of a solar-heated house, described in Example 1.1. We need a model of how the storage temperature $y(t)$ is affected by fan velocity and solar intensity. A straightforward linear model of the type (4.7) would be

$$y(t) + a_1 y(t-1) + a_2 y(t-2) = b_1 u(t-1) + b_2 u(t-2) + c_1 I(t-1) + c_2 I(t-2) \tag{5.13}$$

With this we have not used any physical insight into the heating process, but introduced the black-box model (5.13) in an ad hoc manner. A moment's reflection reveals that a linear model is not very realistic. Clearly, the effects of solar intensity and fan velocity are not additive. When the fan is off, the sun does not at all affect the storage temperature.

Let us go through what happens in the heating system. Introduce $x(t)$ for the temperature of the solar panel collector at time t. With some simplifications, the physics can be described as follows in discrete time: The heating of the air in the collector $[= x(t+1) - x(t)]$ is equal to heat supplied by the sun $[= d_2 \cdot I(t)]$ minus loss of heat to the environment $[= d_3 \cdot x(t)]$ minus the heat transported to storage $[= d_0 \cdot x(t) \cdot u(t)]$; that is,

$$x(t+1) - x(t) = d_2 I(t) - d_3 x(t) - d_0 x(t) \cdot u(t) \tag{5.14}$$

In the same way, the increase of storage temperature $[= y(t+1) - y(t)]$ is equal to supplied heat $[= d_0 x(t) \cdot u(t)]$ minus losses to the environment $[= d_1 y(t)]$; that is,

$$y(t+1) - y(t) = d_0 x(t) u(t) - d_1 y(t) \tag{5.15}$$

In equations (5.14) and (5.15) the coefficients d_k are unknown constants, whose numerical values are to be determined. The temperature $x(t)$ is not, however, measured, so we first eliminate $x(t)$ from (5.14) and (5.15). This gives

$$\begin{aligned} y(t) = {} & (1 - d_1) y(t-1) + (1 - d_3) \frac{y(t-1)u(t-1)}{u(t-2)} \\ & + (d_3 - 1)(1 + d_1) \frac{y(t-2)u(t-1)}{u(t-2)} + d_0 d_2 u(t-1) I(t-2) \\ & - d_0 u(t-1) y(t-1) + d_0 (1 + d_1) u(t-1) y(t-2) \end{aligned} \tag{5.16}$$

The relationship between the measured quantities y, u, and I and the parameters d_i is now more complicated. It can be simplified by reparametrization:

$$\begin{aligned} \theta_1 &= (1 - d_1) & \varphi_1(t) &= y(t-1) \\ \theta_2 &= (1 - d_3) & \varphi_2(t) &= \frac{y(t-1)u(t-1)}{u(t-2)} \\ \theta_3 &= (d_3 - 1)(1 + d_1) & \varphi_3(t) &= \frac{y(t-2)u(t-1)}{u(t-2)} \\ \theta_4 &= d_0 d_2 & \varphi_4(t) &= u(t-1) I(t-2) \\ \theta_5 &= -d_0 & \varphi_5(t) &= u(t-1) y(t-1) \\ \theta_6 &= d_0 (1 + d_1) & \varphi_6(t) &= u(t-1) y(t-2) \\ \theta^T &= [\theta_1 \, \theta_2 \ldots \theta_6] & \varphi^T(t) &= [\varphi_1(t) \, \varphi_2(t) \ldots \varphi_6(t)] \end{aligned} \tag{5.17}$$

Then (5.16) can be rewritten as a true linear regression,

$$y(t) = \hat{y}(t|\theta) = \varphi^T(t)\theta \tag{5.18}$$

where we have a linear relationship between the new parameters θ and the constructed measurements $\varphi(t)$. (Notice that φ does not depend on θ.) The price for this is that the knowledge of algebraic relationships between the θ_i, according to (5.17), has been lost. ■

Hammerstein Model

Sometimes the nonlinearities in a system have the character of a static nonlinearity at the input side, while the dynamics itself is linear, as depicted in Figure 5.1. In case the nonlinearity f is known, we could simply redefine the input as $\bar{u}(t) = f(u(t))$ and treat the system as linear. When f is unknown, it could be approximated by a polynomial expansion

$$f(u) = \alpha_1 u + \alpha_2 u^2 + \cdots + \alpha_m u^m \tag{5.19}$$

and then let each power of u pass a different numerator dynamics:

$$A(q)y(t) = B_1(q)u(t) + B_2(q)u^2(t) + \cdots + B_m(q)u^m(t) \tag{5.20}$$

where $A(q)$ and $B_i(q)$ are polynomials in the delay operator q^{-1}. With

$$\theta^T = [a_1 \ldots a_{n_a} b_1^{(1)} \ldots b_n^{(1)} b_1^{(2)} \ldots b_n^{(2)} \ldots b_1^{(m)} \ldots b_n^{(m)}]$$

$$\varphi^T(t) = [-y(t-1) \ldots -y(t-n_a)\, u(t-1) \ldots$$
$$u(t-n)\, u^2(t-1) \ldots u^2(t-n) \ldots u^m(t-1) \ldots u^m(t-n)]$$

(5.20) can be rewritten

$$y(t) = \hat{y}(t|\theta) = \varphi^T(t)\theta \tag{5.21}$$

and we have a special case of (5.12). The model (5.20) is known as a *Hammerstein model.* It was apparently first discussed in an identification context by Narendra and Gallman (1966).

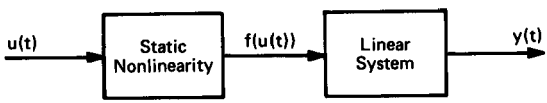

Figure 5.1 A system with a static nonlinearity at the input side.

5.3 NONLINEAR STATE-SPACE MODELS

A General Model Set

The most general description of a finite-dimensional system is

$$\begin{aligned} x(t+1) &= f(t, x(t), u(t), w(t); \theta) \\ y(t) &= h(t, x(t), u(t), v(t); \theta) \end{aligned} \tag{5.22}$$

Here $w(t)$ and $v(t)$ are sequences of independent random variables and θ denotes a vector of unknown parameters. The problem to determine a predictor based on (5.22) and on formal probabilistic grounds is substantial. In fact, this nonlinear prediction problem is known to have no finite-dimensional solution except in some isolated special cases.

Nevertheless, predictors for (5.22) can of course be constructed, either with ad hoc approaches or by some approximation of the unrealizable optimal solution. For

the latter problem, there is abundant literature (see, e.g., Jazwinski, 1970, or Anderson and Moore, 1979). In either case the resulting predictor takes the form

$$\hat{y}(t|\theta) = g(t, Z^{t-1}; \theta) \tag{5.23}$$

Here, for easier notation, we introduced

$$Z^t = (y^t, u^t) = (y(1), u(1) \ldots y(t), u(t))$$

to denote the input–output measurements available at time t. This is the form in which the model is put to use for identification purposes. We may thus view (5.23) as the *basic model,* and disregard the route that took us from the underlying description [like (5.22)] to the form (5.23). This is consistent with the view of *models as predictors* that we took in Chapter 4; the only difference is that (5.23) is a nonlinear function of past data, rather than a linear one. Just as in Chapter 4, the model (5.23) can be complemented with assumptions about the associated prediction error

$$\varepsilon(t, \theta) = y(t) - g(t, Z^{t-1}; \theta) \tag{5.24}$$

such as its covariance matrix, $\Lambda(t; \theta)$ or its PDF $f_e(x, t; \theta)$.

Nonlinear Simulation Model

A particularly simple way of deriving a predictor from (5.22) is to disregard the process noise $w(t)$ and take

$$\begin{aligned} x(t+1, \theta) &= f(t, x(t, \theta), u(t), 0; \theta) \\ \hat{y}(t|\theta) &= h(t, x(t, \theta), u(t), 0; \theta) \end{aligned} \tag{5.25}$$

We call such a predictor a *simulation model,* since $\hat{y}(t|\theta)$ is constructed by simulating a noise-free model (5.22) using the actual input. Clearly, a simulation model is almost as easy to use starting from a continuous-time representation:

$$\begin{aligned} \frac{d}{dt} x(t, \theta) &= f(t, x(t, \theta), u(t), 0; \theta) \\ \hat{y}(t|\theta) &= h(t, x(t, \theta), u(t), 0; \theta) \end{aligned} \tag{5.26}$$

Example 5.2 Delignification

Consider the problem of reducing the lignin content of wood chips in a chemical mixture. This is, basically, the process of cellulose cooking for obtaining pulp for paper making.

Introduce the following notation:

$x(t)$: lignin concentration at time t

$u_1(t)$: absolute temperature at time t

$u_2(t)$: concentration of hydrogen sulfite, $[HSO_3^-]$

$u_3(t)$: concentration of hydrogen, $[H^+]$

Then basic chemical laws tell us that

$$\frac{d}{dt}x(t) = -k_1 e^{-E_L/u_1(t)}[x(t)]^m \cdot [u_2(t)]^\alpha \cdot [u_3(t)]^\beta \tag{5.27}$$

Here E_L is the Arrhenius constant and k_1, m, α, and β are other constants associated with the reaction. Simulating (5.27) with the measured values of $\{u_i(t), i = 1, 2, 3\}$ for given values of $\theta^T = (E_L, k, k_1, m, \alpha, \beta)$ gives a sequence of corresponding lignin concentrations $\{x(t, \theta)\}$. In this case the system output is also the lignin concentration, so $\hat{y}(t|\theta) = x(t, \theta)$. These predicted, or simulated, values can then be compared with the actually measured values so that the errors associated with a particular value of θ can be evaluated. Such an application is described in detail in Hagberg and Schöön (1974). ■

5.4 FORMAL CHARACTERIZATION OF MODELS (*)

In this section we shall give a counterpart of the discussion of Section 4.5 for general, possibly time-varying, and nonlinear models. We assume that the output is p-dimensional and that the input is m-dimensional. Z^t denotes, as before, the input–output data up to and including time t.

Models

A model m of a dynamical system is a sequence of functions $g_m(t, Z^{t-1}), t = 1, 2, \ldots$, from $\mathbf{R} \times \mathbf{R}^{p(t-1)} \times \mathbf{R}^{m(t-1)}$ to $\mathbf{R}^p$, representing a way of guessing or predicting the output $y(t)$ from past data:

$$\hat{y}(t|t-1) = g_m(t, Z^{t-1}) \tag{5.28}$$

A model that defines only the predictor function is called a *predictor model.* When (5.28) is complemented with the conditional (given Z^{t-1}) probability density function (CPDF) of the associated prediction errors

$$f_e(x, t, Z^{t-1}): \quad \text{CPDF of } y(t) - \hat{y}(t|t-1), \qquad \text{given } Z^{t-1} \tag{5.29}$$

we call the model a *complete probabilistic model.* A typical model assumption is that the prediction errors are independent. Then f_e does not depend on Z^{t-1}:

$$f_e(x, t): \quad \text{PDF of } y(t) - \hat{y}(t|t-1): \quad \text{these errors independent} \tag{5.30}$$

Sometimes one may prefer not to specify the complete PDF but only its second moment (the covariance matrix):

$$\Lambda_m(t): \quad \text{covariance matrix of } y(t) - \hat{y}(t|t-1) \quad \text{these errors independent} \tag{5.31}$$

A model (5.28) together with (5.31) could be called a *partial probabilistic model.*

A model can further be classified according to the following properties.

1. The model m is said to be *linear* if $g_m(t, Z^{t-1})$ is linear in y^{t-1} and u^{t-1}:

$$g_m(t, Z^{t-1}) = W_t^y(q)y(t) + W_t^u(q)u(t) \tag{5.32}$$

2. A model m is said to be *time invariant* if $g_m(t, Z^{t-1})$ is invariant under a shift of absolute time. If Λ_m or f_e is specified, it is further required that they be independent of t.
3. A model m is said to be a *k-step-ahead predictor* if $g_m(t, Z^{t-1})$ is a function of y^{t-k}, u^{t-1} only.
4. A model m is said to be a *simulation model* or an *output error model* if $g_m(t, Z^{t-1})$ is a function of u^{t-1} only.

Analogously to the linear case, we could define the stability of the predictor function and equality between different models [see (4.113)]. We refrain, however, from elaborating on these points here.

Model Sets and Model Structures

Sets of models $\mathcal{M}^*$ as well as model structures $\mathcal{M}$ as differentiable mappings

$$\mathcal{M}: \theta \to g(t, Z^{t-1}; \theta) \in \mathcal{M}^*; \theta \in D_{\mathcal{M}} \subset \mathbf{R}^d \tag{5.33}$$

[and $\Lambda(t; \theta)$ or $f_e(x, t; \theta)$ if applicable] from subsets of $\mathbf{R}^d$ to model sets can be defined analogously to Definition 4.3. Once equality between models has been defined, identifiability concepts can be developed as in Section 4.5.

We shall say that a model structure $\mathcal{M}$ is a *linear regression* if $D_{\mathcal{M}} = \mathbf{R}^d$ and the predictor function is a linear (or affine) function of θ:

$$g(t, Z^{t-1}; \theta) = \varphi^T(t, Z^{t-1})\theta + \mu(t, Z^{t-1}) \tag{5.34}$$

Another View of Models (*)

The definition of models as predictors is a rather pragmatic way of approaching the model concept. A more abstract line of thought can be developed as follows.

As users, we communicate with the system only through the input–output data sequences $Z^t = (y^t, u^t)$. Therefore, any assumption about the properties of the system will be an assumption about Z^t. We could thus say that

A model of a system is an assumed relationship for Z^t, $t = 1, 2, \ldots$. (5.35)

Often, experiments on a system are not exactly reproducible. For a given input sequence u^N, we may obtain different output sequences y^N at different experiments due to the presence of various disturbances. In such cases it is natural to regard y^t as a random variable of which we observe different realizations. A model of the system would then be a description of the probabilistic properties of Z^t (or, perhaps, of y^t,

given u^t). This model m could be formulated in terms of a probability measure P_m or the probability density function (PDF) for Z^t:

$$\bar{f}_m(t, Z^t) \tag{5.36}$$

That is,

$$P_m(Z^t \in B) = \int_{x^t \in B} \bar{f}_m(t, x^t)dx^t \tag{5.37}$$

Sometimes it is preferable to consider the input u^t as a given deterministic sequence and focus attention on the conditional PDF of y^t, given u^t:

$$\bar{f}_m(t, y^t|u^t) \tag{5.38}$$

A model (5.36) or (5.38) would normally be quite awkward to construct and work with, and other, indirect ways of forming $\bar{f}_m$ will be preferred. Indeed, the stochastic models of Sections 4.2 and 4.3 are implicit descriptions of the probability density function for the measured signals. The introduction of unmeasurable, stochastic disturbances $\{w(t)\}$, $\{e(t)\}$, and so on, is a convenient way of describing the probabilistic properties of the observed signal, and also often corresponds to an intuitive feeling for how the output is generated. It is, however, worth noting that the effect of these unmeasurable disturbances in the model is just to define the PDF for the observed signals.

The assumed PDF $\bar{f}_m$ in (5.36) is in a sense the most general model that can be applied for an observed data record y^t, u^t. It includes deterministic models as a special case. It also corresponds to a general statistical problem: how to describe the properties of an observed data vector. For our current purposes, it is, however, not a suitably structured model. The natural direction of time flow in the data record, as well as the notions of causality, is not present in (5.36).

Given $\bar{f}_m(t, Z^t)$ in (5.36), it is, at least conceptually, possible to compute the conditional mean of $y(t)$ given y^{t-1}, u^{t-1}; that is,

$$\hat{y}(t|t-1) = E_m[y(t)|y^{t-1}, u^{t-1}] = g_m(t, Z^{t-1}) \tag{5.39}$$

and the distribution of $y(t) - g_m(t, Z^{t-1})$, say $f_e(x, t, Z^{t-1})$. From (5.36) we can thus compute a model (5.28) along with a CPDF f_e in (5.29). Conversely, given the predictor function $g_m(t, Z^{t-1})$ and an assumed PDF $f_e(x, t)$ for the associated prediction errors, we can calculate the joint PDF for the data y^t, u^t as in (5.36). This follows from the following lemma:

Lemma 5.1. Suppose that u^t is a given, deterministic sequence, and assume that the generation of y^t is described by the model

$$y(t) = g_m(t, Z^{t-1}) + \varepsilon_m(t) \tag{5.40}$$

where the conditional PDF of $\varepsilon_m(t)$ (given y^{t-1}, u^{t-1}) is $f_e(x, t)$. Then the joint probability density function for y^t, given u^t, is

$$\bar{f}_m(t, y^t|u^t) = \prod_{k=1}^{t} f_e(y(k) - g_m(k, Z^{k-1}), k) \tag{5.41}$$

Here we have, for convenience, denoted the dummy variable x_k for the distribution of $y(k)$ by $y(k)$ itself.

Proof. The output $y(t)$ is generated by (5.40). Hence the CPDF of $y(t)$, given Z^{t-1}, is

$$p(x_t|Z^{t-1}) = f_e(x_t - g_m(t, Z^{t-1}), t) \tag{5.42}$$

Using Bayes's rule (I.10), the joint CPDF of $y(t)$ and $y(t-1)$, given Z^{t-2}, can be expressed as

$$\begin{aligned} p(x_t, x_{t-1}|Z^{t-2}) &= p(x_t|y(t-1) = x_{t-1}, Z^{t-2}) \cdot p(x_{t-1}|Z^{t-2}) \\ &= f_e(x_t - g_m(t, Z^{t-1}), t) \cdot f_e(x_{t-1} - g_m(t-1, Z^{t-2}), t-1) \end{aligned}$$

where $y(t-1)$ in $g_m(t, Z^{t-1})$ should be replaced by x_{t-1}. Here we have assumed u^t to be a given deterministic sequence. Iterating the preceding expression to $t = 1$ gives the joint probability density function of $y(t), y(t-1), \ldots, y(1)$, given u^t, that is, the function $\bar{f}_m(t, y^t|u^t)$ in (5.41). ■

The important conclusion from this discussion is that the predictor model (5.28), complemented with an assumed PDF for the associated prediction errors, is *no more and no less general* than the general, unstructured joint PDF model (5.36).

Remark. Notice the slight difference, though, in the conditional PDF for the prediction errors. The general form (5.36) may in general lead to a conditional PDF that in fact depends on Z^{t-1}; $f_e(x, t, Z^{t-1})$ as in (5.29). This means that the prediction errors are not necessarily independent, while they do form a martingale difference sequence:

$$E[\varepsilon_m(t)|\varepsilon_m(t-1), \ldots, \varepsilon_m(1)] = 0 \tag{5.43}$$

In the predictor formulation (5.40), we assumed the CPDF $f_e(x, t)$ not to depend on Z^{t-1}, which is an implied assumption of independence of $\varepsilon_m(t)$ on previous data. Clearly, though, we could have relaxed that assumption with obvious modifications in (5.41) as a result. ■

5.5 SUMMARY

The development of models for nonlinear systems is quite analogous to that for linear systems. The basic difference from a formal point of view is that the predictor function becomes a nonlinear function of past observations. The important difference from a practical point of view is that the potential richness of possibilities makes unstructured, "black-box"-type models unfeasible in most cases. Instead, knowledge about the character of the nonlinearities will have to be built into the models. Such structure does not, however, have to be in analytic form. The nonlinearities could very well be defined in look-up tables, and the model parameters could be entries in these tables.

We have also given a short summary of formal aspects of models of dynamical systems. We have stressed that a model in the first place is a predictor function from past observations to the future output. The predictor function may possibly be complemented with a model assumption of properties of the associated prediction error, such as its variance or its PDF.

5.6 BIBLIOGRAPHY

Models for identification of nonlinear systems are discussed in the surveys by Haber and Keviczky (1976), Mehra (1979), and Billings (1980). Leontaritis and Billings (1985) give a thorough treatment of various parametric models, such as NARMAX (nonlinear ARMAX; cf. Problem 5G.1). The Hammerstein model was apparently first discussed in Narendra and Gallman (1966). The nonlinear simulation model, like (5.25) and (5.26), is frequently used in application areas where considerable a priori information is available. A typical such area is identification of aircraft dynamics. See Grübel (1985) for a treatment of such problems.

Applications of various parametric estimation techniques to nonlinear parametric structures are discussed in, for example, Billings and Voon (1984), Gabr and Subba Rao (1984), and Stoica and Söderström (1982b). Recursive techniques are treated by Fnaiech and Ljung (1986).

A general discussion of the model concept is given by Willems (1985).

5.7 PROBLEMS

5G.1. Consider the following nonlinear structure:

$$x(t) = f(x(t-1), \dots, x(t-n), u(t-1), \dots, u(t-n); \theta) \tag{5.44a}$$

$$y(t) = x(t) + v(t) \tag{5.44b}$$

$$v(t) = H(q, \theta)e(t) \tag{5.44c}$$

Here (5.44a) describes the nonlinear noise-free dynamics, parametrized by θ, while (5.44b) describes the measurements as the noise-free output, corrupted by the noise $\{v(t)\}$, which is modeled in the general way (2.19). Show that the natural predictor for (5.44) is given by

$$\hat{y}(t|\theta) = [1 - H^{-1}(q, \theta)]y(t) + H^{-1}(q, \theta)x(t, \theta)$$

where $x(t, \theta)$ is defined by

$$x(t, \theta) = f(x(t-1, \theta), \dots, x(t-n, \theta), u(t-1), \dots, u(t-n); \theta)$$

5E.1. Consider the bilinear model structure described by

$$x(t) + a_1 x(t-1) + a_2 x(t-2) = b_1 u(t-1) + b_2 u(t-2) + c_1 x(t-1)u(t-1)$$

$$y(t) = x(t) + v(t)$$

where

$$\theta = [a_1\, a_2\, b_1\, b_2\, c_1]^T$$

(a) Assume $\{v(t)\}$ to be white noise and compute the predictor $\hat{y}(t|\theta)$ and give an expression for it in the pseudolinear regression form

$$\hat{y}(t|\theta) = \varphi^T(t, \theta)\theta$$

with a suitable vector $\varphi(t, \theta)$.

(b) Now suppose that $\{v(t)\}$ is not white, but can be modeled as an (unknown) first-order ARMA process. Then suggest a suitable predictor for the system.

5E.2. Consider the system in Figure 5.1, where the nonlinearity is saturation with an unknown gain:

$$f(u(t)) = \begin{cases} \theta_1 \cdot \theta_2, & \text{if } u(t) > \theta_2 \\ \theta_1 u(t), & \text{if } |u(t)| < \theta_2 \\ -\theta_1 \cdot \theta_2, & \text{if } u(t) < -\theta_2 \end{cases}$$

Suppose that the linear system can be described by a second-order ARX model. Write down, explicitly, the predictor for this model, parametrized in θ_1, θ_2 and the ARX parameters.

5T.1. Time-continuous bilinear system descriptions are common in many fields (see Mohler, 1973). A model can be written

$$\dot{x}(t) = A(\theta)x(t) + B(\theta)u(t) + G(\theta)x(t) \cdot u(t) + w(t) \tag{5.45a}$$

where $x(t)$ is the state vector, $w(t)$ is white Gaussian noise with variance matrix R_1, and $u(t)$ is a scalar input. The output of the system is sampled as

$$y(t) = C(\theta)x(t) + e(t), \qquad \text{for } t = k \cdot T \tag{5.45b}$$

where $e(t)$ is white Gaussian measurement noise with variance R_2. The input is piecewise constant:

$$u(t) = u_k, \qquad kT \le t < (k+1)T$$

Derive an expression for the prediction of $y((k+1)T)$, given u_r and $y(r \cdot T)$ for $r \le k$, based on the model (5.45).

5T.2 Consider the Monod growth model structure

$$\dot{x}_1 = \frac{\theta_1 \cdot x_2}{\theta_2 + x_2} \cdot x_1 - \alpha_1 x_1$$

$$\dot{x}_2 = \frac{-1}{\theta_3} \cdot \frac{\theta_1 \cdot x_2}{\theta_2 + x_2} \cdot x_1 - \alpha_1(x_2 - \alpha_2)$$

$y = [x_1\ x_2]^T$ is measured and α_1 and α_2 are known constants. Discuss whether the parameters θ_1, θ_2 and θ_3 are identifiable.

Remark: Although we did not give any formal definition of identifiability for nonlinear model structures, they are quite analogous to the definitions in Sections 4.5

and 4.6. Thus, test whether two different parameter values can give the same input-output behavior of the model.

[See Holmberg and Ranta (1982). x_1 here is the concentration of the biomass that is growing, while x_2 is the concentration of the growth limiting substrate. θ_1 is the maximum growth rate, θ_2 is the Michaelis Menten constant, and θ_3 is the yield coefficient.]

II methods

6

NONPARAMETRIC TIME- AND FREQUENCY-DOMAIN METHODS

A linear time-invariant model can be described by its transfer functions or by the corresponding impulse responses, as we found in Chapter 4. In this chapter we shall discuss methods that aim at determining these functions by direct techniques without first selecting a confined set of possible models. Such methods are often also called *nonparametric* since they do not (explicitly) employ a finite-dimensional parameter vector in the search for a best description. We shall discuss the determination of the transfer function $G(q)$ from input to output. Section 6.1 deals with time-domain methods for this, and Sections 6.2 to 6.4 describe frequency-domain techniques of various degrees of sophistication. The determination of $H(q)$ or the disturbance spectrum is discussed in Section 6.5.

It should be noted that throughout this chapter we assume the system to operate in open loop [i.e., $\{u(t)\}$ and $\{v(t)\}$ are independent]. Closed-loop configurations will typically lead to problems for nonparametric methods, as outlined in some of the problems. These issues are discussed in more detail in Chapter 14.

6.1 TRANSIENT-RESPONSE ANALYSIS AND CORRELATION ANALYSIS

Impulse-Response Analysis

If a system that is described by (2.8)

$$y(t) = G_0(q)u(t) + v(t) \tag{6.1}$$

is subjected to a pulse input

$$u(t) = \begin{cases} \alpha, & t = 0 \\ 0, & t \neq 0 \end{cases} \tag{6.2}$$

then the output will be

$$y(t) = \alpha g_0(t) + v(t) \tag{6.3}$$

by definition of G_0 and the impulse response $\{g_0(t)\}$. If the noise level is low, it is thus possible to determine the impulse-response coefficients $\{g_0(t)\}$ from an experiment with a pulse input. The estimates will be

$$\hat{g}(t) = \frac{y(t)}{\alpha} \tag{6.4}$$

and the errors $v(t)/\alpha$. This simple idea is *impulse-response analysis*. Its basic weakness is that many physical processes do not allow pulse inputs of such an amplitude that the error $v(t)/\alpha$ is insignificant compared to the impulse-response coefficients. Moreover, such an input could make the system exhibit nonlinear effects that would disturb the linearized behavior we have set out to model.

Step-Response Analysis

Similarly, a step

$$u(t) = \begin{cases} \alpha, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

applied to (6.1) gives the output

$$y(t) = \alpha \sum_{k=1}^{t} g_0(k) + v(t) \tag{6.5}$$

From this, estimates of $g_0(k)$ could be obtained as

$$\hat{g}(t) = \frac{y(t) - y(t-1)}{\alpha} \tag{6.6}$$

which has an error $[v(t) - v(t-1)]/\alpha$. If we really aim at determining the impulse-response coefficients using (6.6), we would suffer from large errors in most practical applications. However, if the goal is to determine some basic control-related characteristics, such as delay time, static gain, and dominating time constants [i.e., the model (4.47)], step responses (6.5) can very well furnish that information to a sufficient degree of accuracy. In fact, well-known rules for tuning simple regulators such as the Ziegler–Nichols rule (Ziegler and Nichols, 1942) are based on model information reached in step responses.

Based on plots of the step response, some characteristic numbers can be graphically constructed, which in turn can be used to determine parameters in a model of given order. We refer to Rake (1980) for a discussion of such characteristics.

Correlation Analysis

Consider the model description (6.1):

$$y(t) = \sum_{k=1}^{\infty} g_0(k)u(t-k) + v(t) \tag{6.7}$$

If the input is a quasi-stationary sequence [see (2.59)] with

$$\bar{E}\, u(t)u(t-\tau) = R_u(\tau)$$

and

$$\bar{E}\, u(t)v(t-\tau) \equiv 0 \qquad \text{(open-loop operation)}$$

then according to Theorem 2.2 (expressed in the time domain)

$$\bar{E}\, y(t)u(t-\tau) = R_{yu}(\tau) = \sum_{k=1}^{\infty} g_0(k)R_u(k-\tau) \tag{6.8}$$

If the input is chosen as white noise so that

$$R_u(\tau) = \alpha\delta_{\tau 0}$$

then

$$g_0(\tau) = \frac{R_{yu}(\tau)}{\alpha}$$

An estimate of the impulse response is thus obtained from an estimate of $R_{yu}(\tau)$; for example,

$$\hat{R}_{yu}^N(\tau) = \frac{1}{N}\sum_{t=\tau}^{N} y(t)u(t-\tau) \tag{6.9}$$

If the input is not white noise, we may estimate

$$\hat{R}_u^N(\tau) = \frac{1}{N}\sum_{t=\tau}^{N} u(t)u(t-\tau) \tag{6.10}$$

and solve

$$\hat{R}_{yu}^N(\tau) = \sum_{k=1}^{M} \hat{g}(k)\hat{R}_u^N(k-\tau) \tag{6.11}$$

for $\hat{g}(k)$. If the input is open for manipulation, it is of course desirable to choose it so that (6.10) and (6.11) become easy to solve. Equipment for generating such signals and solving for $\hat{g}(k)$ is commercially available. See Godfrey (1980) for a more detailed treatment.

6.2 FREQUENCY-RESPONSE ANALYSIS

Sine-wave Testing

The fundamental physical interpretation of the transfer function $G(z)$ is that the complex number $G(e^{i\omega})$ bears information about what happens to an input sinusoid [see (2.32) to (2.34)]. We thus have for (6.1) that with

$$u(t) = \alpha\cos\omega t, \qquad t = 0, 1, 2, \ldots \tag{6.12}$$

then

$$y(t) = \alpha|G_0(e^{i\omega})|\cos(\omega t + \varphi) + v(t) + \text{transient} \tag{6.13}$$

where

$$\varphi = \arg G_0(e^{i\omega}) \tag{6.14}$$

This property also gives a clue to a simple way of determining $G_0(e^{i\omega})$:

> With the input (6.12), determine the amplitude and the phase shift of the resulting output cosine signal, and calculate an estimate $\hat{G}_N(e^{i\omega})$ based on that information. Repeat for a number of frequencies in the interesting frequency band.

This is known as a frequency analysis and is a simple method for obtaining detailed information about a linear system.

Frequency Analysis by the Correlation Method

With the noise component $v(t)$ present in (6.13), it may be cumbersome to determine $|G_0(e^{i\omega})|$ and φ accurately by graphic methods. Since the interesting component of $y(t)$ is a cosine function of known frequency, it is possible to correlate it out from the noise in the following way. Form the sums

$$I_c(N) = \frac{1}{N}\sum_{t=1}^{N} y(t)\cos\omega t, \qquad I_s(N) = \frac{1}{N}\sum_{t=1}^{N} y(t)\sin\omega t \tag{6.15}$$

Inserting (6.13) into (6.15), ignoring the transient term, gives

$$\begin{aligned} I_c(N) &= \frac{1}{N}\sum_{t=1}^{N} \alpha|G_0(e^{i\omega})|\cos(\omega t + \varphi)\cos\omega t + \frac{1}{N}\sum_{t=1}^{N} v(t)\cos\omega t \\ &= \alpha|G_0(e^{i\omega})|\frac{1}{2}\frac{1}{N}\sum_{t=1}^{N}[\cos\varphi + \cos(2\omega t + \varphi)] \\ &\quad + \frac{1}{N}\sum_{t=1}^{N} v(t)\cos\omega t \\ &= \frac{\alpha}{2}|G_0(e^{i\omega})|\cos\varphi + \alpha|G_0(e^{i\omega})|\frac{1}{2}\frac{1}{N}\sum_{t=1}^{N}\cos(2\omega t + \varphi) \\ &\quad + \frac{1}{N}\sum_{t=1}^{N} v(t)\cos\omega t \end{aligned} \tag{6.16}$$

The second term tends to zero as N tends to infinity, and so does the third term if $v(t)$ does not contain a pure periodic component of frequency ω. If $\{v(t)\}$ is a stationary stochastic process such that

$$\sum_0^\infty \tau |R_v(\tau)| < \infty$$

then the variance of the third term of (6.16) decays like $1/N$ (Problem 6T.2). Similarly,

$$I_s(N) = -\frac{\alpha}{2}|G_0(e^{i\omega})|\sin\varphi + \alpha|G_0(e^{i\omega})|\frac{1}{2}\frac{1}{N}\sum_{t=1}^{N}\sin(2\omega t + \varphi) + \frac{1}{N}\sum_{t=1}^{N} v(t)\sin\omega t \tag{6.17}$$

These two expressions suggest the following estimates of $|G_0(e^{i\omega})|$ and φ:

$$|\hat{G}_N(e^{i\omega})| = \frac{\sqrt{I_c^2(N) + I_s^2(N)}}{\alpha/2} \tag{6.18a}$$

$$\hat{\varphi}_N = \arg \hat{G}_N(e^{i\omega}) = -\arctan\frac{I_s(N)}{I_c(N)} \tag{6.18b}$$

Rake (1980) and Åström (1975) give a more detailed account of this method. By repeating the procedure for a number of frequencies, a good picture of $G_0(e^{i\omega})$ over the frequency domain of interest can be obtained. Equipment that performs such *frequency analysis by the correlation method* is commercially available.

An advantage with this method is that a Bode plot of the system can be obtained easily and that one may concentrate the effort to the interesting frequency ranges. The main disadvantage is that many industrial processes do not admit sinusodial inputs in normal operation. The experiment must also be repeated for a number of frequencies which may lead to long experimentation periods.

Relationship to Fourier Analysis

Comparing (6.15) to the definition (2.37),

$$Y_N(\omega) = \frac{1}{\sqrt{N}}\sum_{t=1}^{N} y(t)e^{-i\omega t} \tag{6.19}$$

shows that

$$I_c(N) - iI_s(N) = \frac{1}{\sqrt{N}}Y_N(\omega) \tag{6.20}$$

As in (2.46) we find that, for (6.12),

$$U_N(\xi) = \frac{\sqrt{N}\,\alpha}{2}, \qquad \text{if } \xi = \omega = \frac{2\pi r}{N} \text{ for some integer } r \tag{6.21}$$

It is straightforward to rearrange (6.18) as

$$\hat{G}_N(e^{i\omega}) = \frac{\sqrt{N}\,Y_N(\omega)}{N\alpha/2} \tag{6.22}$$

which, using (6.21), means that

$$\hat{G}_N(e^{i\omega}) = \frac{Y_N(\omega)}{U_N(\omega)} \tag{6.23}$$

Here ω is precisely the frequency of the input signal. Comparing with (2.53), we also find (6.23) a most reasonable estimate [especially since $R_N(\omega)$ in (2.53) is zero for periodic inputs, according to the corollary of Theorem 2.1].

6.3 FOURIER ANALYSIS

Empirical Transfer-function Estimate

We found the expression (6.23) to correspond to frequency analysis with a single sinusoid of frequency ω as input. In a linear system, different frequencies pass through the system independently of each other. It is therefore quite natural to extend the frequency analysis estimate (6.23) also to the case of multifrequency inputs. That is, we introduce the following estimate of the transfer function:

$$\hat{\hat{G}}_N(e^{i\omega}) = \frac{Y_N(\omega)}{U_N(\omega)} \tag{6.24}$$

with Y_N and U_N defined by (6.19) and (2.37), respectively, also for the case where the input is not a single sinusoid. This estimate is also quite natural in view of Theorem 2.1.

We shall call $\hat{\hat{G}}_N(e^{i\omega})$ the *empirical transfer-function estimate* (ETFE), for reasons that we shall discuss shortly. In (6.24) we assume of course that $U_N(\omega) \neq 0$. If this does not hold for some frequencies, we simply regard the ETFE as undefined at those frequencies. We call this estimate empirical, since no other assumptions have been imposed than linearity of the system. In the case of multifrequency inputs, the ETFE consists of $N/2$ essential points. [Recall that estimates at frequencies intermediate to the grid $\omega = 2\pi k/N, k = 0, 1, \ldots, N-1$, are obtained by trigonometrical interpolation in (2.37). Also, since y and u are real, we have

$$\hat{\hat{G}}_N(e^{2\pi ik/N}) = \overline{\hat{\hat{G}}_N(e^{2\pi i(N-k)/N})} \tag{6.25}$$

(compare (2.40) and (2.41)].

The original data sequence consisting of $2N$ numbers $y(t), u(t), t = 1, 2, \ldots, N$, has thus been condensed into the N numbers

$$\operatorname{Re} \hat{\hat{G}}_N(e^{2\pi ik/N}), \qquad \operatorname{Im} \hat{\hat{G}}_N(e^{2\pi ik/N}), \qquad k = 0, 1, \ldots, \frac{N}{2} - 1$$

This is quite a modest data reduction, revealing that most of the information contained in the original data y, u still is quite "raw."

In addition to an extension of frequency analysis, the ETFE can be interpreted as a way of (approximately) solving the set of convolution equations

$$y(t) = \sum_{k=1}^{N} g_0(k)u(t-k), \qquad t = 1, 2, \ldots, N \tag{6.26}$$

for $g_0(k), k = 1, 2, \ldots, N$, using Fourier techniques.

Properties of the ETFE

Assume that the system is subject to (6.1). Introducing

$$V_N(\omega) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} v(t)e^{-i\omega t} \tag{6.27}$$

for the disturbance term, we find from Theorem 2.1 that

$$\hat{\hat{G}}_N(e^{i\omega}) = G_0(e^{i\omega}) + \frac{R_N(\omega)}{U_N(\omega)} + \frac{V_N(\omega)}{U_N(\omega)} \tag{6.28}$$

where the term $R_N(\omega)$ is subject to (2.54) and decays as $1/\sqrt{N}$.

Let us now investigate the influence of the term $V_N(\omega)$ on $\hat{\hat{G}}_N(e^{i\omega})$. Since $v(t)$ is assumed to have zero mean value,

$$EV_N(\omega) = 0, \qquad \forall\omega$$

so that

$$E\hat{\hat{G}}_N(e^{i\omega}) = G_0(e^{i\omega}) + \frac{R_N(\omega)}{U_N(\omega)} \tag{6.29}$$

Here expectation is with respect to $\{v(t)\}$, assuming $\{u(t)\}$ to be a given sequence of numbers.

Let the covariance function $R_v(\tau)$ and the spectrum $\Phi_v(\omega)$ of the process $\{v(t)\}$ be defined by (2.14) and (2.63). Then evaluate

$$\begin{aligned} EV_N(\omega)V_N(-\xi) &= \frac{1}{N}\sum_{r=1}^{N}\sum_{s=1}^{N} Ev(r)e^{-i\omega r}v(s)e^{+i\xi s} \\ &= \frac{1}{N}\sum_{r=1}^{N}\sum_{s=1}^{N} e^{i(\xi-\omega)r}R_v(r-s) = [r-s=\tau] \\ &= \frac{1}{N}\sum_{r=1}^{N} e^{i(\xi-\omega)r}\cdot\sum_{\tau=r-N}^{r-1} R_v(\tau)e^{-i\xi\tau} \end{aligned}$$

Now

$$\sum_{\tau=r-N}^{r-1} R_v(\tau)e^{-i\xi\tau} = \Phi_v(\xi) - \sum_{\tau=-\infty}^{r-N-1} e^{-i\xi\tau}R_v(\tau) - \sum_{\tau=r}^{\infty} e^{-i\xi\tau}R_v(\tau)$$

and

$$\frac{1}{N}\sum_{r=1}^{N} e^{i(\xi-\omega)r} = \begin{cases} 1, & \text{if } \xi = \omega \\ 0, & \text{if } (\xi-\omega) = \dfrac{k2\pi}{N}, \end{cases} \qquad k = \pm 1, \pm 2, \ldots, \pm(N-1)$$

Consider

$$\left|\frac{1}{N}\sum_{r=1}^{N} e^{i(\xi-\omega)r}\cdot\sum_{\tau=-\infty}^{r-N-1} e^{-i\xi\tau}R_v(\tau)\right| \le \frac{1}{N}\sum_{r=1}^{N}\sum_{\tau=-\infty}^{r-N-1}|R_v(\tau)|$$

$$\le [\text{change order of summation}]$$

$$\le \frac{1}{N}\sum_{\tau=-\infty}^{-1}|\tau|\cdot|R_v(\tau)| \le \frac{C}{N}$$

provided

$$\sum_{-\infty}^{\infty}|\tau\cdot R_v(\tau)| < \infty \tag{6.30}$$

Similarly,

$$\left|\frac{1}{N}\sum_{r=1}^{N} e^{i(\xi-\omega)r}\cdot\sum_{\tau=r}^{\infty} e^{-i\xi\tau}R_v(\tau)\right| \le \frac{1}{N}\sum_{\tau=1}^{\infty}\tau\cdot|R_v(\tau)| \le \frac{C}{N}$$

Combining these expressions, we find that

$$EV_N(\omega)V_N(-\xi) = \begin{cases}\Phi_v(\omega)+\rho_2(N), & \text{if } \xi=\omega\\ \rho_2(N), & \text{if } |\xi-\omega|=\dfrac{k2\pi}{N}, \quad k=1,2,\ldots,N-1\end{cases} \tag{6.31}$$

with

$$|\rho_2(N)| \le \frac{2C}{N}$$

These calculations can be summarized as the following result.

Lemma 6.1. Consider a strictly stable system

$$y(t) = G_0(q)u(t) + v(t) \tag{6.32}$$

with a disturbance $\{v(t)\}$ being a stationary stochastic process with spectrum $\Phi_v(\omega)$ and covariance function $R_v(\tau)$, subject to (6.30). Let $\{u(t)\}$ be independent of $\{v(t)\}$ and assume that $|u(t)| \le C$ for all t. Then with $\hat{\hat{G}}_N(e^{i\omega})$ defined by (6.24), we have

$$E\hat{\hat{G}}_N(e^{i\omega}) = G_0(e^{i\omega}) + \frac{\rho_1(N)}{U_N(\omega)} \tag{6.33a}$$

where

$$|\rho_1(N)| \le \frac{C_1}{\sqrt{N}} \tag{6.33b}$$

and

$$E[\hat{\hat{G}}_N(e^{i\omega}) - G_0(e^{i\omega})][\hat{\hat{G}}_N(e^{-i\xi}) - G_0(e^{-i\xi})]$$

$$= \begin{cases} \dfrac{1}{|U_N(\omega)|^2}[\Phi_v(\omega) + \rho_2(N)], & \text{if } \xi = \omega \\ \dfrac{\rho_2(N)}{U_N(\omega)U_N(-\xi)}, & \text{if } |\xi - \omega| = \dfrac{2\pi k}{N}, \quad k = 1, 2, \ldots, N-1 \end{cases} \tag{6.34a}$$

where

$$|\rho_2(N)| \le \frac{C_2}{N} \tag{6.34b}$$

Here U_N is defined by (2.37), and we restrict ourselves to frequencies for which $\hat{\hat{G}}_N$ is defined. According to Theorem 2.1 and (6.30), the constants can be taken as

$$C_1 = \left(2\sum_{k=1}^{\infty}|kg_0(k)|\right)\cdot \max|u(t)| \tag{6.35a}$$

$$C_2 = C_1^2 + \sum_{k=-\infty}^{\infty}|\tau R_v(\tau)| \tag{6.35b}$$

If $\{u(t)\}$ is periodic with period N, then $C_1 = 0$. ■

Remark. Note that the input is regarded as a given sequence. Probabilistic quantities, such as E, "bias," and "variance" refer to the probability space of $\{v(t)\}$. This does not, of course, exclude that the input may be generated as a realization of a stochastic process independent of $\{v(t)\}$. ■

The properties of the ETFE are closely related to those of periodogram estimates of spectra. See (2.43) and (2.72). We have the following result.

Lemma 6.2. Let $v(t)$ be given by

$$v(t) = H(q)e(t)$$

where $\{e(t)\}$ is a white-noise sequence with variance λ and fourth moment μ^2, and H is a strictly stable filter. Let $V_N(\omega)$ be defined by (6.27), and let $\Phi_v(\omega)$ be the spectrum of $v(t)$. Then

$$E|V_N(\omega)|^2 = \Phi_v(\omega) + \rho_3(N) \tag{6.36}$$

$$E(|V_N(\omega)|^2 - \Phi_v(\omega))(|V_N(\xi)|^2 - \Phi_v(\xi))$$

$$= \begin{cases} [\Phi_v(\omega)]^2 + \rho_4(N), & \text{if } \xi = \omega \quad \omega \ne 0, \pi \\ \rho_4(N) & \text{if } |\xi - \omega| = \dfrac{2\pi k}{N}, \quad k = 1, 2, \ldots, N-1 \end{cases} \tag{6.37}$$

where

$$|\rho_3(N)| \le \frac{C}{N}, \qquad |\rho_4(N)| \le \frac{C}{N}$$

Proof. Equation (6.36) is a restatement of (6.31). A simple proof of (6.37) is outlined in Problem 6D.2 under somewhat more restrictive conditions. A full proof can be given by direct evaluation of (6.37). See, for example, Brillinger (1981), Theorem 5.2.4, for that. See Problem 6G.5 for ideas on how the bias term can be improved by the use of data tapering. ■

These lemmas, together with the results of Section 2.3, tell us the following:

Case 1. The input is periodic. When the input is periodic and N is a multiple of the period, we know from Example 2.2 that $|U_N(\omega)|^2$ increases like const $\cdot N$ for some ω and is zero for others [see (2.49)]. The number of frequencies $\omega = 2\pi k/N$ for which $|U_N(\omega)|^2$ is nonzero, and hence for which the ETFE is defined, is fixed and no more than the period length of the signal. We thus find that

- The ETFE $\hat{G}_N(e^{i\omega})$ is defined only for a fixed number of frequencies.
- At these frequencies the ETFE is unbiased and its variance decays like $1/N$.

We note that the results (6.17) on frequency analysis by the correlation method are obtained as a special case.

Case 2. The input is a realization of stochastic process. Lemma 6.2 shows that the periodogram $|U_N(\omega)|^2$ is an erratic function of ω, which fluctuates around $\Phi_u(\omega)$, which we assume to be bounded. Lemma 6.1 thus tells us that

- The ETFE is an asymptotically unbiased estimate of the transfer function at increasingly (with N) many frequencies.
- The variance of the ETFE does not decrease as N increases, and it is given as the noise-to-signal ratio at the frequency in question.
- The estimates at different frequencies are asymptotically uncorrelated.

It follows from this discussion that, in the case of a periodic input signal, the ETFE will be of increasingly good quality at the frequencies that are present in the input. However, when the input is not periodic, the variance does not decay with N, but remains equal to the noise-to-signal ratio at the corresponding frequency. This latter property makes the empirical estimate a very crude estimate in most cases in practice.

It is easy to understand the reason why the variance does not decrease with N. We determine as many independent estimates as we have data points. In other words, we have no feature of data and information compression. This in turn is due to the fact that we have only assumed linearity about the true system. Consequently, the system's properties at different frequencies may be totally unrelated. From this it also follows that the only possibility to increase the information per estimated parameter is to assume that the system's behavior at one frequency is related to that

at another. In the subsequent section, we shall discuss one approach to how this can be done.

6.4. SPECTRAL ANALYSIS

Spectral analysis for determining transfer functions of linear systems was developed from statistical methods for spectral estimation. Good accounts of this method are given in Chapter 10 in Jenkins and Watts (1968) and in Chapter 6 in Brillinger (1981), and the method is widely discussed in many other textbooks on time series analysis. In this section we shall adopt a slightly unusual approach to the subject by deriving the standard techniques as a smoothed version of the ETFE.

Smoothing the ETFE

We mentioned at the end of the previous section that the only way to improve on the poor variance properties of the ETFE is to assume that the values of the true transfer function at different frequencies are related. We shall now introduce the rather reasonable prejudice that

The true transfer function $G_0(e^{i\omega})$ is a smooth function of ω. (6.38)

If now the frequency distance $2\pi/N$ is small compared to how quickly $G_0(e^{i\omega})$ changes, then

$$\hat{\hat{G}}_N(e^{2\pi ik/N}), \qquad k \text{ integer}, \qquad 2\pi k/N \approx \omega \tag{6.39}$$

are uncorrelated, unbiased estimates of roughly the same constant $G_0(e^{i\omega})$, each with a variance of

$$\frac{\Phi_v(2\pi k/N)}{|U_N(2\pi k/N)|^2}$$

according to Lemma 6.1. Here we neglected terms that tend to zero as N tends to infinity.

If we assume $G_0(e^{i\omega})$ to be constant over the interval

$$\frac{2\pi k_1}{N} = \omega_0 - \Delta\omega < \omega < \omega_0 + \Delta\omega = \frac{2\pi k_2}{N} \tag{6.40}$$

then it is well known that the best (in a minimum variance sense) way to estimate this constant is to form a weighted average of the "measurements" (6.39) for the frequencies (6.40), each measurement weighted according to its inverse variance [compare Problem 6E.3, and Lemma II.2, (II.58), in Appendix II]:

$$\hat{G}_N(e^{i\omega_0}) = \frac{\sum_{k=k_1}^{k_2} \alpha_k \hat{\hat{G}}_N(e^{2\pi ik/N})}{\sum_{k=k_1}^{k_2} \alpha_k} \tag{6.41a}$$

$$\alpha_k = \frac{|U_N(2\pi k/N)|^2}{\Phi_v(2\pi k/N)} \tag{6.41b}$$

For large N we could with good approximation work with the integrals that correspond to the (Riemann) sums in (6.41a):

$$\hat{G}_N(e^{i\omega_0}) = \frac{\int_{\xi=\omega_0-\Delta\omega}^{\omega_0+\Delta\omega} \alpha(\xi)\hat{\hat{G}}_N(e^{i\xi})d\xi}{\int_{\xi=\omega_0-\Delta\omega}^{\omega_0+\Delta\omega} \alpha(\xi)d\xi} \tag{6.42a}$$

$$\alpha(\xi) = \frac{|U_N(\xi)|^2}{\Phi_v(\xi)} \tag{6.42b}$$

If the transfer function G_0 is not constant over the interval (6.40) it is reasonable to use an additional weighting that pays more attention to frequencies close to ω_0:

$$\hat{G}_N(e^{i\omega_0}) = \frac{\int_{-\pi}^{\pi} W_\gamma(\xi-\omega_0)\alpha(\xi)\hat{\hat{G}}_N(e^{i\xi})d\xi}{\int_{-\pi}^{\pi} W_\gamma(\xi-\omega_0)\alpha(\xi)d\xi} \tag{6.43}$$

Here $W_\gamma(\xi)$ is a function centered around $\xi = 0$ and γ is a "shape parameter," which we shall discuss shortly.

Clearly, (6.42) corresponds to

$$W_\gamma(\xi) = \begin{cases} 1, & |\xi| < \Delta\omega \\ 0, & |\xi| > \Delta\omega \end{cases} \tag{6.44}$$

Now, if the noise spectrum $\Phi_v(\omega)$ is known, the estimate (6.43) can be realized as written. If $\Phi_v(\omega)$ is not known we could argue as follows: Suppose that the noise spectrum does not change very much over frequency intervals corresponding to the "width" of the weighting function $W_\gamma(\xi)$:

$$\int_{-\pi}^{\pi} W_\gamma(\xi-\omega_0)\left|\frac{1}{\Phi_v(\xi)} - \frac{1}{\Phi_v(\omega_0)}\right| d\xi = \text{"small"} \tag{6.45}$$

Then $\alpha(\xi)$ in (6.42b) can be replaced by $\alpha(\xi) = |U_N(\xi)|^2/\Phi_v(\omega_0)$, which means that the constant $\Phi_v(\omega_0)$ cancels when (6.43) is formed. Under (6.45) the estimate

$$\hat{G}_N(e^{i\omega_0}) = \frac{\int_{-\pi}^{\pi} W_\gamma(\xi-\omega_0)|U_N(\xi)|^2\hat{\hat{G}}_N(e^{i\xi})\,d\xi}{\int_{-\pi}^{\pi} W_\gamma(\xi-\omega_0)|U_N(\xi)|^2\,d\xi} \tag{6.46}$$

is thus a good approximation of (6.42b) and (6.43).

We may remark that, if (6.45) does not hold, it might be better to include a procedure where $\Phi_v(\omega)$ is estimated and use that estimate in (6.43).

Connection with the Blackman–Tukey Procedure (*)

Consider the denominator of (6.46). It is a weighted average of the periodogram $|U_N(\xi)|^2$. Using the result (2.72), we find that, as $N \to \infty$,

$$\int_{-\pi}^{\pi} W_\gamma(\xi - \omega_0)|U_N(\xi)|^2 d\xi \to \int_{-\pi}^{\pi} W_\gamma(\xi - \omega_0)\Phi_u(\xi)d\xi \tag{6.47}$$

where $\Phi_u(\omega)$ is the spectrum of $\{u(t)\}$, as defined by (2.61) to (2.63). If, moreover,

$$\int_{-\pi}^{\pi} W_\gamma(\xi)d\xi = 1$$

and the weighting function $W_\gamma(\xi)$ is concentrated around $\xi = 0$ with a width over which $\Phi_u(\omega)$ does not change much, then the right side of (6.47) is close to $\Phi_u(\omega_0)$. We may thus interpret the left side as an estimate of this quantity:

$$\hat{\Phi}_u^N(\omega_0) = \int_{-\pi}^{\pi} W_\gamma(\xi - \omega_0)|U_N(\xi)|^2 d\xi \tag{6.48}$$

Similarly, since

$$|U_N(\xi)|^2 \hat{\hat{G}}_N(e^{i\xi}) = |U_N(\xi)|^2 \frac{Y_N(\xi)}{U_N(\xi)} = Y_N(\xi)\overline{U}_N(\xi) \tag{6.49}$$

we have that the numerator of (6.46)

$$\hat{\Phi}_{yu}^N(\omega_0) = \int_{-\pi}^{\pi} W_\gamma(\xi - \omega_0)Y_N(\xi)\overline{U}_N(\xi)d\xi \tag{6.50}$$

is an estimate of the cross spectrum between output and input. The transfer function estimate (6.46) is thus the ratio of two spectral estimates:

$$\hat{G}_N(e^{i\omega_0}) = \frac{\hat{\Phi}_{yu}^N(\omega_0)}{\hat{\Phi}_u^N(\omega_0)} \tag{6.51}$$

which makes sense, in view of (2.78). The spectral estimates (6.48) and (6.50) are the standard estimates, suggested in the literature, for spectra and cross spectra as smoothed periodograms. See Blackman and Tukey (1958), Jenkins and Watts (1968), or Brillinger (1981).

An alternative way of expressing these estimates is common. The Fourier coefficients for the periodogram $|U_N(\omega)|^2$ are

$$\hat{R}_u^N(\tau) = \frac{1}{2\pi}\int_{-\pi}^{\pi} |U_N(\omega)|^2 e^{i\tau\omega} d\omega = \frac{1}{N}\sum_{t=1}^{N} u(t)u(t - \tau) \tag{6.52}$$

[For this expression to hold exactly, the values $u(s)$ outside the interval $1 \le s \le N$ have to be interpreted by periodic continuation; i.e., $u(s) = u(s - N)$ if $s > N$; see Problem 6D.1.]

Similarly, let the Fourier coefficients of the function $W_\gamma(\xi)$ be

$$w_\gamma(\tau) = \int_{-\pi}^{\pi} W_\gamma(\xi)\, e^{i\xi\tau}\, d\xi \tag{6.53}$$

Since the integral (6.48) is a convolution, its Fourier coefficients will be the product of (6.52) and (6.53), so a Fourier expansion of (6.48) gives

$$\hat{\Phi}_u^N(\omega) = \sum_{\tau=-\infty}^{\infty} w_\gamma(\tau) \cdot \hat{R}_u^N(\tau) e^{-i\tau\omega} \tag{6.54}$$

The idea is now that the nice, smooth function $W_\gamma(\xi)$ is chosen so that its Fourier coefficients vanish for $|\tau| > \delta_\gamma$, where typically $\delta_\gamma << N$. It is consequently sufficient to form (6.52) (using the rightmost expression) for $|\tau| \le \delta_\gamma$ and then take

$$\hat{\Phi}_u^N(\omega) = \sum_{\tau=-\delta_\gamma}^{\delta_\gamma} w_\gamma(\tau) \hat{R}_u^N(\tau) e^{-i\tau\omega} \tag{6.55}$$

This is perhaps the most convenient way of forming the spectral estimate. The expressions for $\hat{\Phi}_{yu}^N(\omega)$ are of course analogous.

Weighting Function $W_\gamma(\xi)$: The Frequency Window

Let us now discuss the weighting function $W_\gamma(\xi)$. In spectral analysis, it is often called the *frequency window*. [Similarly, $w_\gamma(\tau)$ is called the *lag window*.] If this window is "wide," then many different frequencies will be weighted together in (6.40). This should lead to a small variance of $\hat{G}_N(e^{i\omega_0})$. At the same time, a wide window will involve frequency estimates farther away from ω_0, with expected values that may differ considerably from $G_0(e^{i\omega_0})$. This will cause large bias. The width of the window will thus control the trade-off between bias and variance. To make this trade-off a bit more formal, we shall use the scalar γ to describe the width, so a large value of γ corresponds to a narrow window.

We shall characterize the window by the following numbers

$$\int_{-\pi}^{\pi} W_\gamma(\xi) d\xi = 1, \qquad \int_{-\pi}^{\pi} \xi W_\gamma(\xi) d\xi = 0, \qquad \int_{-\pi}^{\pi} \xi^2 W_\gamma(\xi) d\xi = M(\gamma) \tag{6.56a}$$

$$\int_{-\pi}^{\pi} |\xi|^3 W_\gamma(\xi) d\xi = C_3(\gamma), \qquad \int_{-\pi}^{\pi} W_\gamma^2(\xi) d\xi = \frac{1}{2\pi} \overline{W}(\gamma) \tag{6.56b}$$

As γ increases (and the frequency window gets more narrow), the number $M(\gamma)$ decreases, while $\overline{W}(\gamma)$ increases.

Some typical windows are given in Table 6.1. [See, also, Table 3.3.1 in Brillinger (1981) for a more complete collection of windows.] Notice that the scaling

TABLE 6.1 Some Windows for Spectral Analysis

	$2\pi \cdot W_\gamma(\omega)$	$w_\gamma(\tau), \quad 0 \le \lvert\tau\rvert \le \gamma$
Bartlett	$\frac{1}{\gamma}\left(\frac{\sin \gamma\omega/2}{\sin \omega/2}\right)^2$	$1 - \frac{\lvert\tau\rvert}{\gamma}$
Parzen	$\frac{4(2+\cos\omega)}{\gamma^3}\left(\frac{\sin \gamma\omega/4}{\sin \omega/2}\right)^4$	$1 - \frac{6\tau^2}{\gamma^2}\left(1 - \frac{\lvert\tau\rvert}{\gamma}\right), \quad 0 \le \lvert\tau\rvert \le \frac{\gamma}{2}$ $2\left(1 - \frac{\lvert\tau\rvert}{\gamma}\right)^3, \quad \frac{\gamma}{2} \le \lvert\tau\rvert \le \gamma$
Hamming	$\frac{1}{2}D_\gamma(\omega) + \frac{1}{4}D_\gamma\left(\omega - \frac{\pi}{\gamma}\right) + \frac{1}{4}D_\gamma\left(\omega + \frac{\pi}{\gamma}\right)$, where $D_\gamma(\omega) = \frac{\sin(\gamma + \frac{1}{2})\omega}{\sin \omega/2}$	$\frac{1}{2}\left(1 + \cos\frac{\pi\tau}{\gamma}\right)$

quantity γ has been chosen so that $\delta_\gamma = \gamma$ in (6.55). The frequency windows are shown graphically in Figure 6.1. For these windows, we have

$$\begin{aligned} \text{Bartlett:} \quad & M(\gamma) \approx \frac{2.78}{\gamma}, \quad && \overline{W}(\gamma) \approx 0.67\gamma \\ \text{Parzen:} \quad & M(\gamma) \approx \frac{12}{\gamma^2}, && \overline{W}(\gamma) \approx 0.54\gamma \\ \text{Hamming:} \quad & M(\gamma) \approx \frac{\pi^2}{2\gamma^2}, && \overline{W}(\gamma) \approx 0.75\gamma \end{aligned} \tag{6.57}$$

The expressions are asymptotic for large γ but are good approximations for $\gamma \gtrsim 5$. See also Problem 6T.1 for a further discussion of how to scale windows.

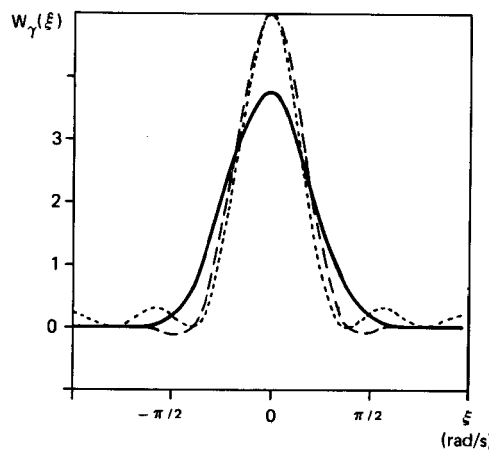

Figure 6.1 Some common frequency windows. Solid line: Parzen; dashed line: Hamming; dotted line: Bartlett, $\gamma = 5$.

Asymptotic Properties of the Smoothed Estimate

The estimate (6.46) has been studied in several treatments of spectral analysis. See, for example, Chapter 10 of Jenkins and Watts (1968) or Chapter 6 of Brillinger (1981). Results that are asymptotic in both N and γ can be derived as follows (see Appendix 6A). Consider the estimate (6.46), and suppose that the true system obeys the assumptions of Lemma 6.1. We then have

Bias

$$E\hat{G}_N(e^{i\omega}) - G_0(e^{i\omega}) = M(\gamma) \cdot \left[\frac{1}{2} G_0''(e^{i\omega}) + G_0'(e^{i\omega}) \frac{\Phi_u'(\omega)}{\Phi_u(\omega)} \right] + \underset{\gamma \to \infty}{\mathbf{O}}(C_3(\gamma)) + \underset{N \to \infty}{\mathbf{O}}(1/\sqrt{N}) \tag{6.58}$$

Prime and double prime denote differentiation with respect to ω, once and twice, respectively.

Variance

$$E|\hat{G}_N(e^{i\omega}) - E\hat{G}_N(e^{i\omega})|^2 = \frac{1}{N} \cdot \overline{W}(\gamma) \cdot \frac{\Phi_v(\omega)}{\Phi_u(\omega)} + \underset{\substack{\gamma \to \infty \\ N \to \infty, \gamma/N \to 0}}{\mathbf{o}}(\overline{W}(\gamma)/N) \tag{6.59}$$

We repeat that expectation here is with respect to the noise sequence $\{v(t)\}$ and that the input is supposed to be a deterministic quasi-stationary signal.

Let us use the asymptotic expressions to evaluate the mean-square error (MSE):

$$E|\hat{G}_N(e^{i\omega}) - G_0(e^{i\omega})|^2 \sim M^2(\gamma)|R(\omega)|^2 + \frac{1}{N}\overline{W}(\gamma)\frac{\Phi_v(\omega)}{\Phi_u(\omega)} \tag{6.60}$$

Here

$$R(\omega) = \frac{1}{2} G_0''(e^{i\omega}) + G_0'(e^{i\omega}) \frac{\Phi_u'(\omega)}{\Phi_u(\omega)} \tag{6.61}$$

Some additional results can also be shown (see Brillinger, 1981, Chapter 6, and Problems 6D.3 and 6D.4).

- The estimates $\operatorname{Re} \hat{G}_N(e^{i\omega})$ and $\operatorname{Im} \hat{G}_N(e^{i\omega})$ are asymptotically uncorrelated and each have a variance equal to half that in (6.59). (6.62)
- The estimates $\hat{G}_N(e^{i\omega})$ at different frequencies are asymptotically uncorrelated. (6.63)
- The estimates $\operatorname{Re} \hat{G}_N(e^{i\omega_k})$, $\operatorname{Im} \hat{G}_N(e^{i\omega_k})$, $k = 1, 2, \ldots, M$, at an arbitrary collection of frequencies are asymptotically jointly normal distributed with means and covariances given by (6.58) to (6.63). (6.64)
- For a translation to properties of $|\hat{G}_N(e^{i\omega})|$, $\arg \hat{G}_N(e^{i\omega})$, see Problem 9G.1.

From (6.60) we see that a desired property of the window is that both M and $\overline{W}$ should be small. We may also calculate the value of the width parameter γ that

minimizes the MSE. Suppose that both γ and N tend to infinity and γ/N tends to zero, so that the asymptotic expressions are applicable. Suppose also that (6.57) holds with $M(\gamma) = M/\gamma^2$ and $\overline{W}(\gamma) = \gamma \cdot \overline{W}$ Then (6.60) gives

$$\gamma_{\text{opt}} = \left(\frac{4M^2 |R(\omega)|^2 \Phi_u(\omega)}{\overline{W} \Phi_v(\omega)} \right)^{1/5} \cdot N^{1/5} \tag{6.65}$$

This value can of course not be realized by the user, since the constant contains several unknown quantities. We note, however, that in any case it increases like $N^{1/5}$, and it should, in principle, be allowed to be frequency dependent. The frequency window consequently should get more narrow when more data are available, which is a very natural result.

The optimal choice of γ leads to a mean-square error that decays like

$$\text{MSE} \sim C \cdot N^{-4/5} \tag{6.66}$$

In practical use the trade-off (6.65) and (6.66) cannot be reached in formal terms. Instead, a typical procedure would be to start by taking $\gamma = N/20$ (see Table 6.1) and then compute and plot the corresponding estimates $\hat{G}_N(e^{i\omega})$ for various values of γ. As γ is increased, more and more details of the estimate will appear. These will be due to decreased bias (true resonance peaks appearing more clearly and the like), as well as to increased variance (spurious, random peaks). The procedure will be stopped when the user feels that the emerging details are predominately spurious.

An Example

Example 6.1

The system

$$y(t) - 1.5y(t-1) + 0.7y(t-2) = u(t-1) + 0.5u(t-2) + e(t) \tag{6.67}$$

where $\{e(t)\}$ is white noise with variance 1 was simulated with the input as a PRBS signal (see Section 14.3) over 1000 samples. Part of the resulting data record is shown in Figure 6.2. The corresponding ETFE is shown in Figure 6.3. An estimate $\hat{G}_N(e^{i\omega})$ was formed using (6.46), with $W_\gamma(\xi)$ being a Parzen window with various values of γ. Figure 6.4 shows the results for $\gamma = 10$, 50, and 200. Here $\gamma = 50$ appears to be a reasonable choice of window size. ■

Another Way of Smoothing the ETFE (*)

The guiding idea behind the estimate (6.46) is that the ETFEs at neighboring frequencies are asymptotically uncorrelated, and that hence the variance could be reduced by averaging over these. The ETFEs obtained over different data sets will also provide uncorrelated estimates, and another approach would be to form averages over these. Thus, split the data set Z^N into M batches, each containing R data ($N = R \cdot M$). Then form the ETFE corresponding to the kth batch:

$$\hat{G}_R^{(k)}(e^{i\omega}), \qquad k = 1, 2, \ldots, M \tag{6.68}$$

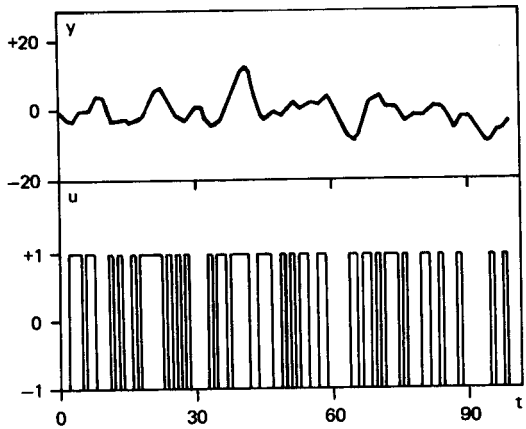

Figure 6.2 The simulated data from (6.67).

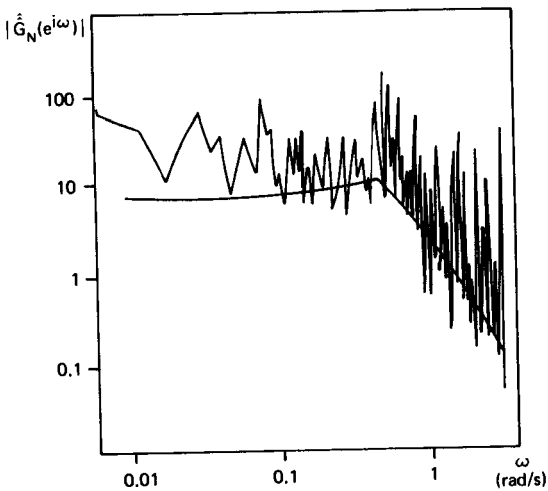

Figure 6.3 The amplitude plot of ETFE based on the data in Figure 6.2. Smooth line: true amplitude plot of (6.67).

The estimate can then be formed as a direct average

$$\hat{G}_N(e^{i\omega}) = \frac{1}{M}\sum_{k=1}^{M} \hat{G}_R^{(k)}(e^{i\omega}) \tag{6.69}$$

or one that is weighted according to the inverse variances:

$$\hat{G}_N(e^{i\omega}) = \frac{\sum_{k=1}^{M} \beta_R^{(k)}(\omega)\cdot \hat{G}_R^{(k)}(e^{i\omega})}{\sum_{k=1}^{M} \beta_R^{(k)}(\omega)} \tag{6.70}$$

with

$$\beta_R^{(k)}(\omega) = |U_R^{(k)}(\omega)|^2 \tag{6.71}$$

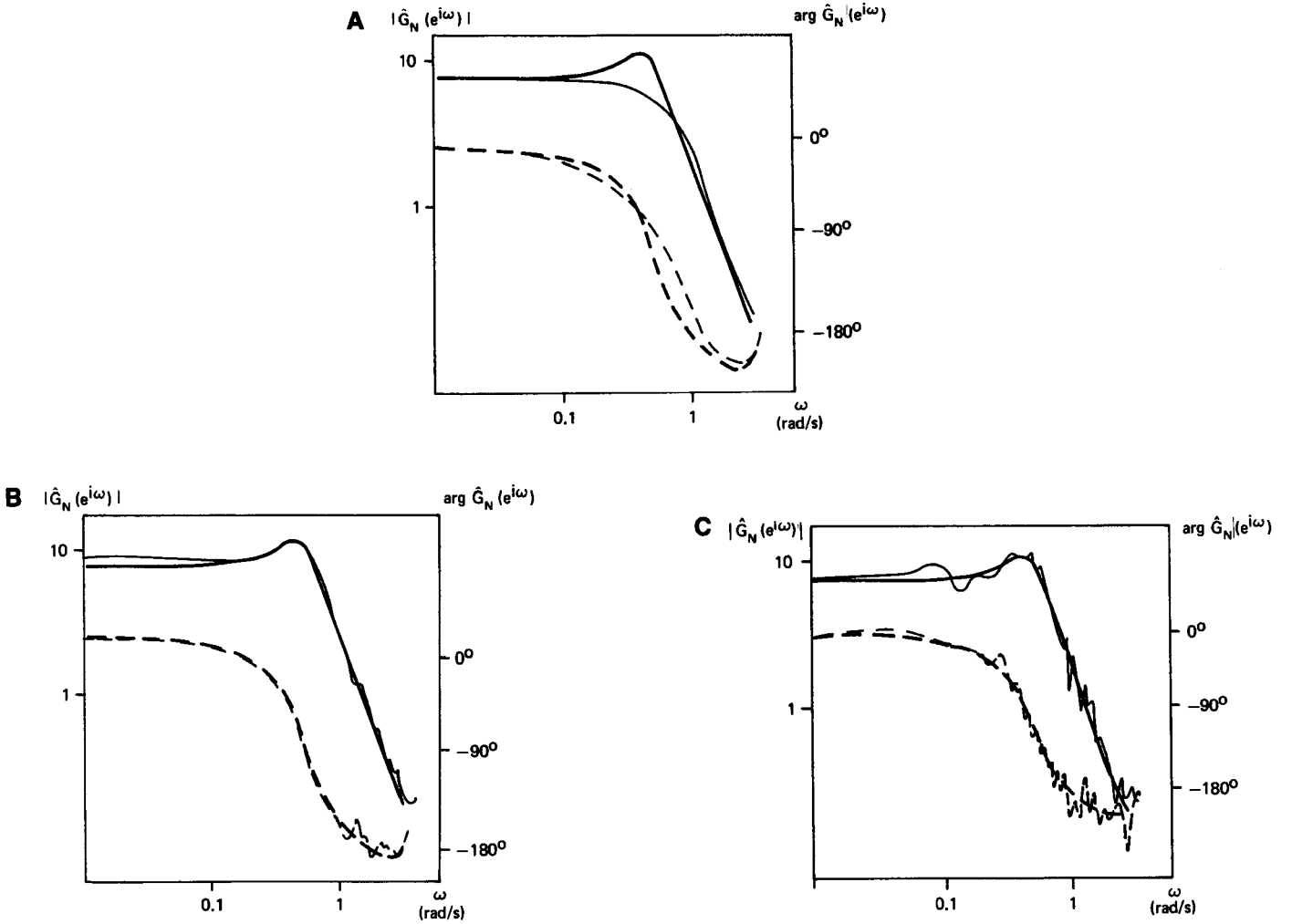

Figure 6.4 The estimate $\hat{G}_N(e^{i\omega})$ for $\gamma = 10$ (A), 50 (B) and 200 (C). Solid lines: amplitude plots; dashed lines: phase plots; thick lines: true system; thin lines: estimate.

being the periodogram of the kth subbatch. The inverse variance of $\hat{\bar{G}}_R^{(k)}(e^{i\omega})$ is $\beta_R^{(k)}(\omega)/\Phi_v(\omega)$, but the factor $\Phi_v(\omega)$ cancels when (6.70) is formed.

An advantage with the estimate (6.70) is that the fast Fourier transform (FFT) can be efficiently used when Z^N can be decomposed so that R is a power of 2. Compare Problem 6G.4.

6.5 ESTIMATING THE DISTURBANCE SPECTRUM (*)

Estimating Spectra

So far we have described how to estimate G_0 in a relationship (6.1):

$$y(t) = G_0(q)u(t) + v(t) \tag{6.72}$$

We shall now turn to the problem of estimating the spectrum of $\{v(t)\}$, $\Phi_v(\omega)$. Had the disturbances $v(t)$ been available for direct measurement, we could have used (6.48):

$$\hat{\Phi}_v^N(\omega) = \int_{-\pi}^{\pi} W_\gamma(\xi - \omega)|V_N(\xi)|^2 \, d\xi \tag{6.73}$$

Here $W_\gamma(\cdot)$ is a frequency window of the kind described earlier.

It is entirely analogous to the analysis of the previous section to calculate the properties of (6.73). We have:

Bias

$$E\hat{\Phi}_v^N(\omega) - \Phi_v(\omega) = \frac{1}{2}M(\gamma)\cdot\Phi_v''(\omega) + \mathbf{O}\underset{\gamma\to\infty}{(C_1(\gamma))} + \mathbf{O}\underset{N\to\infty}{(1/\sqrt{N})} \tag{6.74}$$

Variance

$$\operatorname{Var}\hat{\Phi}_v^N(\omega) = \frac{\overline{W}(\gamma)}{N}\cdot\Phi_v^2(\omega) + \mathbf{o}\underset{N\to\infty}{(1/N)}, \qquad \omega \neq 0, \pm\pi \tag{6.75}$$

Moreover, estimates at different frequencies are asymptotically uncorrelated.

The Residual Spectrum

Now the $v(t)$ in (6.72) are not directly measurable. However, given an estimate $\hat{G}_N$ of the transfer function, we may replace v in the preceding expression by

$$\hat{v}(t) = y(t) - \hat{G}_N(q)u(t) \tag{6.76}$$

which gives the estimate

$$\hat{\Phi}_v^N(\omega) = \int_{-\pi}^{\pi} W_\gamma(\xi - \omega)|Y_N(\xi) - \hat{G}_N(e^{i\xi})U_N(\xi)|^2 \, d\xi \tag{6.77}$$

If $\hat{G}_N(e^{i\xi})$ is formed using (6.46) with the same window $W_\gamma(\cdot)$, this expression can be rearranged as follows [using (6.48) to (6.51)]:

$$\begin{aligned}
&\int_{-\pi}^{\pi} W_\gamma(\xi-\omega)|Y_N(\xi)|^2\,d\xi + \int_{-\pi}^{\pi} W_\gamma(\xi-\omega)|U_N(\xi)|^2|\hat{G}_N(e^{i\xi})|^2\,d\xi \\
&\quad - 2\mathrm{Re}\int_{-\pi}^{\pi} W_\gamma(\xi-\omega)\hat{G}_N(e^{i\xi})U_N(\xi)\overline{Y_N(\xi)}\,d\xi \\
&\approx \int_{-\pi}^{\pi} W_\gamma(\xi-\omega)|Y_N(\xi)|^2\,d\xi + |\hat{G}_N(e^{i\omega})|^2\int_{-\pi}^{\pi} W_\gamma(\xi-\omega)|U_N(\xi)|^2\,d\xi \\
&\quad - 2\mathrm{Re}\,\hat{G}_N(e^{i\omega})\int_{-\pi}^{\pi} W_\gamma(\xi-\omega)U_N(\xi)\overline{Y_N(\xi)}\,d\xi \\
&= \hat{\Phi}_y^N(\omega) + \frac{|\hat{\Phi}_{yu}^N(\omega)|^2}{(\hat{\Phi}_u^N(\omega))^2}\cdot\hat{\Phi}_u^N(\omega) - 2\mathrm{Re}\frac{\hat{\Phi}_{yu}^N(\omega)}{\hat{\Phi}_u^N(\omega)}\cdot\overline{\hat{\Phi}_{yu}^N(\omega)}
\end{aligned}$$

Here the approximate equality follows from replacing the smooth function $\hat{G}_N(e^{i\xi})$ over the small interval around $\xi = \omega$ with its value at ω. Hence we have

$$\hat{\Phi}_v^N(\omega) = \hat{\Phi}_y^N(\omega) - \frac{|\hat{\Phi}_{yu}^N(\omega)|^2}{\hat{\Phi}_u^N(\omega)} \tag{6.78}$$

Asymptotically, as $N\to\infty$ and $\gamma\to\infty$, so that $\hat{G}_N(e^{i\omega})\to G_0(e^{i\omega})$ according to (6.60), we find that the estimate (6.77) tends to (6.73). The asymptotic properties (6.74) and (6.75) will also hold for (6.77) and (6.78). In addition to the properties already listed, we may note that the estimates $\hat{\Phi}_v^N(\omega)$ are asymptotically uncorrelated with $\hat{G}_N(e^{i\omega})$. Moreover, $\hat{\Phi}_v^N(\omega_k), \hat{G}_N(e^{i\omega_k}), k = 1, 2, \ldots, r$, are asymptotically jointly normal random variables with mean and covariances given by (6.58) to (6.64) and (6.74) to (6.75). A detailed account of the asymptotic theory is given in Chapter 6 of Brillinger (1981).

Coherency Spectrum

Denote

$$\hat{\kappa}_{yu}^N(\omega) = \sqrt{\frac{|\hat{\Phi}_{yu}^N(\omega)|^2}{\hat{\Phi}_y^N(\omega)\hat{\Phi}_u^N(\omega)}} \tag{6.79}$$

Then

$$\hat{\Phi}_v^N(\omega) = \hat{\Phi}_y^N(\omega)[1 - (\hat{\kappa}_{yu}^N(\omega))^2] \tag{6.80}$$

The function $\kappa_{yu}(\omega)$ is called the *coherency spectrum* (between y and u) and can be viewed as the (frequency dependent) correlation coefficient between the input and output sequences. If this coefficient is 1 at a certain frequency, then there is perfect

correlation between input and output at that frequency. There is consequently no noise interfering at that frequency, which is confirmed by (6.80).

6.6 SUMMARY

In this chapter we have shown how simple techniques of transient and frequency response can give valuable insight into the properties of linear systems. We have introduced the empirical transfer-function estimate (ETFE)

$$\hat{\hat{G}}_N(e^{i\omega}) = \frac{Y_N(\omega)}{U_N(\omega)} \tag{6.81}$$

based on data over the interval $1 \le t \le N$. Here

$$Y_N(\omega) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} y(t)e^{-it\omega}, \qquad U_N(\omega) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} u(t)e^{-it\omega}$$

The ETFE has the property (see Lemma 6.1) that it is asymptotically unbiased, but has a variance of $\Phi_v(\omega)/|U_N(\omega)|^2$.

We showed how smoothing the ETFE leads to the spectral analysis estimate

$$\hat{G}_N(e^{i\omega}) = \frac{\int_{-\pi}^{\pi} W_\gamma(\xi - \omega)|U_N(\xi)|^2 \hat{\hat{G}}_N(e^{i\xi})d\xi}{\int_{-\pi}^{\pi} W_\gamma(\xi - \omega)|U_N(\xi)|^2 d\xi} \tag{6.82}$$

A corresponding estimate of the noise spectrum is

$$\hat{\Phi}_v^N(\omega) = \int_{-\pi}^{\pi} W_\gamma(\xi - \omega)|Y_N(\xi) - \hat{G}_N(e^{i\xi})U_N(\xi)|^2 d\xi \tag{6.83}$$

The properties of these estimates were summarized in (6.58) to (6.64) and (6.74) to (6.75).

These properties depend on the parameter γ, which describes the width of the associated frequency window W_γ. A narrow such window (large γ) gives small bias but high variance for the estimate, while the converse is true for wide windows.

6.7 BIBLIOGRAPHY

Section 6.1: Wellstead (1981) gives a general survey of nonparametric methods for system identification. A survey of transient response methods is given in Rake (1980). Several ways of determining numerical characteristics from step responses are discussed in Schwarze (1964). Correlation techniques are surveyed in Godfrey (1980).

Section 6.2: Frequency analysis is a classical identification method that is described in many textbooks on control. For detailed treatments, see Rake (1980) and Åström (1975), which also contains several interesting examples.

Section 6.3: General Fourier techniques are also discussed in Rake (1980). The term "empirical transfer function estimate" for $\hat{G}$ is introduced in this chapter, but the estimate as such is well known.

Sections 6.4 and 6.5: Spectral analysis is a standard subject in textbooks on time series. See, for example, Grenander and Rosenblatt (1957) (Chapters 4 to 6), T. W. Anderson (1971) (Chapter 9), and Hannan (1970) (Chapter V). These texts deal primarily with estimation of power (auto-) spectra. Among specific treatments of frequency-domain techniques, including estimation of transfer functions, we note Brillinger (1981) for a thorough analytic study, Jenkins and Watts (1968) for a more leisurely discussion of both statistical properties and application aspects, and Bendat and Piersol (1980) for an application-oriented approach. Another extensive treatment is Priestley (1981). Overviews of different frequency-domain techniques are given in Brillinger and Krishnaiah (1983), and a control-oriented survey is given by Godfrey (1980). The treatment given here is based on Ljung (1985a). The first reference to the idea of smoothing the periodogram to obtain a better spectral estimate appears to be Daniell (1946). A comparative discussion of windows for spectral analysis is given in Gecklini and Yavuz (1978) and Papoulis (1973).

Estimation of transfer functions by spectral analysis has been widely used: in e.g. econometrics (Granger, 1964); geophysics (Robinson, 1967); industrial applications (Gustavsson, 1975); and many other areas. The literature is abundant with examples.

In addition to direct frequency-domain methods for estimating spectra, many efficient methods are based on parametric fit, such as those to be discussed in the following chapter. So called maximum entropy methods (MEM) have found wide use in signal-processing applications. See Burg (1967) for the first idea and Kay and Marple (1981) for a comparative survey of different approaches.

6.8 PROBLEMS

6G.1. Consider the system

$$y(t) = G_0(q)u(t) + v(t)$$

controlled by the regulator

$$u(t) = -F(q)y(t) + r(t)$$

where $r(t)$ is an external reference signal. r and v are independent and their spectra are $\Phi_r(\omega)$ and $\Phi_v(\omega)$, respectively. The usual spectral analysis estimate of G_0 is given in

(6.51) as well as (6.46). Show that as N and γ tend to infinity then $\hat{G}_N(e^{i\omega})$ will converge to

$$G_*(e^{i\omega}) = \frac{G_0(e^{i\omega})\Phi_r(\omega) - F(e^{-i\omega})\Phi_v(\omega)}{\Phi_r(\omega) + |F(e^{i\omega})|^2\Phi_v(\omega)}$$

What happens in the two special cases $\Phi_r \equiv 0$ and $F \equiv 0$, respectively? *Hint:* Compare Problem 2E.5.

6G.2. *Prefiltering.* Prefilter inputs and outputs:

$$u_F(t) = L_u(q)u(t), \qquad y_F(t) = L_y(q)y(t)$$

If (6.32) holds, then the filtered variables obey

$$y_F(t) = G_0^F(q)u_F(t) + v_F(t)$$

$$G_0^F(q) = \frac{L_y(q)}{L_u(q)}G_0(q), \qquad v_F(t) = L_y(q)v(t)$$

Apply spectral analysis to u_F, y_F, thus forming an estimate $\hat{G}_N^F(e^{i\omega})$. The estimate of the original transfer function then is

$$\hat{G}_N(e^{i\omega}) = \frac{L_u(e^{i\omega})}{L_y(e^{i\omega})}\hat{G}_N^F(e^{i\omega})$$

Determine the asymptotic properties of $\hat{G}_N(e^{i\omega})$ and discuss how L_u and L_y can be chosen for smallest MSE (cf. Ljung, 1985a).

6G.3. In Figure 6.3 the amplitude of the ETFE appears to be systematically larger than the true amplitude, despite the fact that the ETFE is unbiased according to Lemma 6.1. However, $\hat{G}$ being an unbiased estimate of G_0 does not imply that $|\hat{G}|$ is an unbiased estimate of $|G_0|$. In fact, prove that

$$E|\hat{\hat{G}}_N(e^{i\omega})|^2 = |G_0(e^{i\omega})|^2 + \frac{\Phi_v(\omega)}{|U_N(\omega)|^2}$$

asymptotically for large N, under the assumptions of Lemma 6.1.

6G.4. The Cooley–Tukey spectral estimate for a process $\{v(t)\}$ is defined as

$$\hat{\Phi}_v^N(\omega) = \frac{1}{M}\sum_{k=1}^{M}|V_R^{(k)}(\omega)|^2$$

where $|V_R^{(k)}(\omega)|^2$ is the periodogram estimate of the kth subbatch of data:

$$V_R^{(k)}(\omega) = \frac{1}{\sqrt{R}}\sum_{\ell=1}^{R} v((k-1)\cdot R + \ell)e^{-i\omega\ell}$$

See Cooley and Tukey (1965) or Hannan (1970), Chapter V. The cross-spectral estimate is defined analogously. This estimate has the advantage that the FFT (fast Fourier transform) can be applied (most efficiently if R is a power of 2). Show that the estimate (6.70) is the ratio of two appropriate Cooley–Tukey spectral estimates.

6G.5. *"Tapers" or "faders."* The bias term $\rho_3(N)$ in (6.36) can be reduced if tapering is introduced: Let $V_N^{(T)}(\omega)$ be defined by

$$V_N^{(T)}(\omega) = \sum_{t=1}^{N} h_t v(t)\cdot e^{-it\omega}$$

where $\{h_t\}_1^N$ is a sequence of numbers (a *tapering function*) such that

$$\sum_{t=1}^{N} h_t^2 = 1$$

Let

$$H_N(\omega) = \sum_{t=1}^{N} h_t e^{-it\omega}$$

Show that, under the conditions of Lemma 6.2,

$$E|V_N^{(T)}(\omega)|^2 = \int_{-\pi}^{\pi} |H_N(\omega - \xi)|^2 \Phi_v(\xi) d\xi$$

Show that our standard periodogram estimate, which uses $h_t \equiv 1/\sqrt{N}$, gives

$$|H_N(\omega)|^2 = \frac{1}{N}\left[\frac{\sin N\omega/2}{\sin \omega/2}\right]^2$$

Other tapering coefficients (or "faders" or "convergence factors") may give functions $|H_N(\omega)|^2$ that are more "δ-function like" than in preceding equation (see, e.g., Table 6.1). The tapered periodogram can of course also be used to obtain smoothed spectra. They will typically lead to decreased bias and (slightly) increased variance (Brillinger, 1981, Theorem 5.2.3 and Section 5.8).

6E.1. Determine an estimate for $G_0(e^{i\omega})$ based on the impulse-response estimates (6.4). Show that this estimate coincides with the ETFE (6.24).

6E.2. Consider the system

$$y(t) = G_0(q)u(t) + v(t)$$

This system is controlled by output proportional feedback

$$u(t) = -Ky(t)$$

Let the ETFE $\hat{\hat{G}}_N(e^{i\omega})$ be computed in the straightforward way (6.24). What will this estimate be? Compare with Lemma 6.1.

6E.3. Let $w_k, k = 1, \ldots, M$, be independent random variables, all with mean value 1 and variances $E(w_k - 1)^2 = \lambda_k$. Consider

$$w = \sum_{k=1}^{M} \alpha_k w_k$$

Determine $\alpha_k, k = 1, \ldots, M$, so that

(a) $Ew = 1$.

(b) $E(w - 1)^2$ is minimized.

6T.1. A general approach to treat the relationships between the scaling parameter γ and the lag and frequency windows $w_\gamma(\tau)$ and $W_\gamma(\omega)$ [see (6.53)] can be given as follows. Choose an even function $w(x)$ such that $w(0) = 1$ and $w(x) = 0, |x| > 1$, with Fourier transform

$$W(\lambda) = \int_{-\infty}^{\infty} w(x) e^{-ix\lambda} dx$$

Let

$$\overline{W} = 2\pi \int_{-\infty}^{\infty} W^2(\lambda) d\lambda$$

$$M = \int_{-\infty}^{\infty} \lambda^2 W(\lambda) d\lambda$$

Then define the lag window

$$w_\gamma(\tau) = w(\tau/\gamma)$$

This gives a frequency window

$$W_\gamma(\omega) = \sum_{\tau=-\gamma}^{\gamma} w_\gamma(\tau) e^{-i\tau\omega}$$

Show that, for large γ,

(a) $W_\gamma(\omega) \approx \gamma \cdot W(\gamma \cdot \omega)$

(b) $M(\gamma) \approx M/\gamma^2$

(c) $\overline{W}(\gamma) \approx \overline{W} \cdot \gamma$

where $M(\gamma)$ and $\overline{W}(\gamma)$ are defined by (6.56). Moreover, compute and compare $W_\gamma(\omega)$ and $\gamma \cdot W(\gamma \cdot \omega)$ for $w(x) = 1 - |x|, |x| \le 1$ (the Bartlett window). [Compare (6.57). See also Hannan (1970), Section V.4.]

6T.2. Let $\{v(t)\}$ be a stationary stochastic process with zero mean value and covariance function $R_v(\tau)$, such that

$$\sum_{-\infty}^{\infty} |\tau R_v(\tau)| < \infty$$

Let

$$S_N = \frac{1}{N} \sum_{t=1}^{N} \alpha_t v(t), \qquad |\alpha_t| \le C_1$$

Show that

$$E|S_N|^2 \le \frac{C_2}{N}$$

for some constant C_2.

6D.1. Prove (6.52) with the proper interpretation of values outside the interval $1 \le t \le N$.

6D.2. Prove a relaxed version of Lemma 6.2 with $|\rho_k(N)| \le C/\sqrt{N}$ by a direct application of Theorem 2.1 and the properties of periodograms of white noise.

6D.3. Prove (6.63) by using expressions analogous to (6A.3) and (6A.4).

6D.4. Prove (6.62) by using (6.63) and

$$\operatorname{Re} \hat{G}(e^{i\omega}) = \frac{\hat{G}(e^{i\omega}) + \hat{G}(e^{-i\omega})}{2}$$

$$\operatorname{Im} \hat{G}(e^{i\omega}) = \frac{\hat{G}(e^{i\omega}) - \hat{G}(e^{-i\omega})}{2i}$$

6S.1. Write a MACRO

$$[\text{GH}, \text{PHIVH}] = \text{SPA}(y, u, \text{GAMMA})$$

that implements (6.46) and (6.77). Pick a particular window in Table 6.1 with $\gamma = \text{GAMMA}$. The easiest implementation is to use (6.9) and (6.10) for $|\tau| \leq \gamma$, then (6.55) [and its counterparts for $\hat{\Phi}_y^N(\omega)$, $\hat{\Phi}_{yu}^N(\omega)$], and finally (6.51) and (6.78). Return GH and PHIVH computed at predetermined frequencies as vectors.

APPENDIX 6A: DERIVATION OF THE ASYMPTOTIC PROPERTIES OF THE SPECTRAL ANALYSIS ESTIMATE

Consider the transfer function estimate (6.46). In this appendix we shall derive the asymptotic properties (6.58) and (6.59). In order not to get too technical, some elements of the derivation will be kept heuristic. Recall that $\{u(t)\}$ here is regarded as a deterministic quasi-stationary sequence, and, hence, such that (6.47) holds.

We then have

$$\begin{aligned} E\hat{G}_N(e^{i\omega_0}) &= \frac{\int_{-\pi}^{\pi} W_\gamma(\xi - \omega_0)|U_N(\xi)|^2[G_0(e^{i\xi}) + \rho_1(N)/U_N(\xi)]\,d\xi}{\int_{-\pi}^{\pi} W_\gamma(\xi - \omega_0)|U_N(\xi)|^2\,d\xi} \\ &\approx \frac{\int_{-\pi}^{\pi} W_\gamma(\xi - \omega_0)\Phi_u(\xi)G_0(e^{i\xi})\,d\xi}{\int_{-\pi}^{\pi} W_\gamma(\xi - \omega_0)\Phi_u(\xi)\,d\xi} \end{aligned} \tag{6A.1}$$

using first Lemma 6.1 and then (6.47), neglecting the decaying term $\rho_1(N)$.

Now, expanding in Taylor series (prime denoting differentiation with respect to ω),

$$G_0(e^{i\xi}) \approx G_0(e^{i\omega_0}) + (\xi - \omega_0)G_0'(e^{i\omega_0}) + \tfrac{1}{2}(\xi - \omega_0)^2 G_0''(e^{i\omega_0})$$

$$\Phi_u(\xi) \approx \Phi_u(\omega_0) + (\xi - \omega_0)\Phi_u'(\omega_0) + \tfrac{1}{2}(\xi - \omega_0)^2 \Phi_u''(\omega_0)$$

and noting that, according to (6.56a),

$$\int_{-\pi}^{\pi} (\xi - \omega_0) \cdot W_\gamma(\xi - \omega_0)d\xi = 0$$

$$\int_{-\pi}^{\pi} (\xi - \omega_0)^2 \cdot W_\gamma(\xi - \omega_0)d\xi = M(\gamma)$$

we find that the numerator of (6A.1) is approximately

$$G_0(e^{i\omega_0})\Phi_u(\omega_0) + M(\gamma)[\tfrac{1}{2}\Phi_u'' G_0 + \tfrac{1}{2}G_0''\Phi_u + \Phi_u' G_0']$$

and the denominator

$$\Phi_u(\omega_0) + \tfrac{1}{2}M(\gamma)[\Phi_u'']$$

where we neglect effects that are of order $C_3(\gamma)$ [an order of magnitude smaller than $M(\gamma)$ as $\gamma \to \infty$; see (6.56b)]. Equation (6A.1) thus gives

$$E\hat{G}_N(e^{i\omega_0}) \approx G_0(e^{i\omega_0}) + M(\gamma)\left[\tfrac{1}{2}G_0''(e^{i\omega_0}) + G_0'(e^{i\omega_0})\frac{\Phi_u'(\omega_0)}{\Phi_u(\omega_0)}\right]$$

which is (6.58).

For the variance expression, we first have from (6.28) that

$$\hat{G}_N(e^{i\omega_0}) - E\hat{G}_N(e^{i\omega_0}) \approx \frac{\int_{-\pi}^{\pi} W_\gamma(\xi - \omega_0)|U_N(\xi)|^2[V_N(\xi)/U_N(\xi)]d\xi}{\int_{-\pi}^{\pi} W_\gamma(\xi - \omega_0)|U_N(\xi)|^2\, d\xi} \tag{6A.2}$$

Let us study the numerator of this expression. We write this, approximately, as a Riemann sum [see (6.41a); we could have kept it discrete all along]:

$$\int_{-\pi}^{\pi} W_\gamma(\xi - \omega_0)\overline{U}_N(\xi)V_N(\xi)d\xi \approx A_N$$

$$\triangleq \frac{2\pi}{N}\sum_{k=-(N/2)+1}^{N/2} W_\gamma\left(\frac{2\pi k}{N} - \omega_0\right)\overline{U}_N\left(\frac{2\pi k}{N}\right)V_N\left(\frac{2\pi k}{N}\right) \tag{6A.3}$$

We have, with summation from $1 - N/2$ to $N/2$,

$$\begin{aligned} EA_N\overline{A}_N &= \frac{4\pi^2}{N^2}\sum_k\sum_\ell W_\gamma\left(\frac{2\pi k}{N} - \omega_0\right)W_\gamma\left(\frac{2\pi \ell}{N} - \omega_0\right)\overline{U}_N\left(\frac{2\pi k}{N}\right) \\ &\quad \times \overline{U}_N\left(-\frac{2\pi\ell}{N}\right)EV_N\left(\frac{2\pi k}{N}\right)V_N\left(-\frac{2\pi\ell}{N}\right) \qquad (6A.4) \\ &\approx \frac{4\pi^2}{N^2}\sum\left[W_\gamma\left(\frac{2\pi k}{N} - \omega_0\right)\right]^2 \cdot \left|U_N\left(\frac{2\pi k}{N}\right)\right|^2 \Phi_v\left(\frac{2\pi k}{N}\right) \end{aligned}$$

using (6.31) and neglecting the term $\rho_2(N)$.

Returning to the integral form, we thus have, using (6.47)

$$EA_N\overline{A}_N \approx \frac{2\pi}{N}\int_{-\pi}^{\pi} W_\gamma^2(\xi - \omega_0)\Phi_u(\xi)\Phi_v(\xi)d\xi \approx \frac{1}{N}\overline{W}(\gamma)\Phi_u(\omega_0)\Phi_v(\omega_0)$$

using (6.56) and the fact that, for large γ, $W_\gamma(\xi)$ is concentrated around $\xi = 0$.

The denominator of (6A.2) approximately equals $\Phi_u(\omega_0)$ for the same reason. We thus find that

$$\mathrm{Var}[\hat{G}_N(e^{i\omega_0})] \approx \frac{(1/N)\overline{W}(\gamma)\Phi_u(\omega_0)\Phi_v(\omega_0)}{[\Phi_u(\omega_0)]^2}$$

and (6.59) has been established. ■

7

PARAMETER ESTIMATION METHODS

Suppose a set of candidate models has been selected, and it is parametrized as a model structure (see Section 4.5), using a parameter vector θ. The search for the best model within the set then becomes a problem of determining or estimating θ. There are many different ways of organizing such a search and also different views on what one should search for. In the present chapter we shall concentrate on the latter aspect: what should be meant by a "good model"? Computational issues (i.e., how to organize the actual search) will be dealt with in Chapters 10 and 11. The evaluation of the properties of the models that result under various conditions and using different methods is carried out in Chapters 8 and 9. In Chapter 15 we return to the estimation methods, and give a more user-oriented summary of recommended procedures.

7.1 GUIDING PRINCIPLES BEHIND PARAMETER ESTIMATION METHODS

Parameter Estimation Methods

We are now in the situation that we have selected a certain model structure $\mathcal{M}$, with particular models $\mathcal{M}(\theta)$ parametrized using the parameter vector $\theta \in D_{\mathcal{M}} \subset \mathbf{R}^d$. The set of models thus defined consequently is

$$\mathcal{M}^* = \{\mathcal{M}(\theta) | \theta \in D_{\mathcal{M}}\} \tag{7.1}$$

Recall that each model represents a way of predicting future outputs. The predictor could be a linear filter, as discussed in Chapter 4:

$$\mathcal{M}(\theta): \quad \hat{y}(t|\theta) = W_y(q,\theta)y(t) + W_u(q,\theta)u(t) \tag{7.2}$$

This could correspond to one-step-ahead prediction for an underlying system description

$$y(t) = G(q,\theta)u(t) + H(q,\theta)e(t) \tag{7.3}$$

in which case

$$W_y(q,\theta) = [1 - H^{-1}(q,\theta)], \qquad W_u(q,\theta) = H^{-1}(q,\theta)G(q,\theta) \tag{7.4}$$

but it could also be arrived at from other considerations.

The predictor could also be a nonlinear filter, as discussed in Chapter 5, in which case we write it as a general function of past data Z^{t-1}:

$$\mathcal{M}(\theta): \quad \hat{y}(t|\theta) = g(t, Z^{t-1}; \theta) \tag{7.5}$$

The model $\mathcal{M}(\theta)$ may also contain (model) assumptions about the character of the associated prediction errors, such as their variances ($\lambda(\theta)$) or their probability distribution (PDF $f_e(x, \theta)$).

We are also in the situation that we have collected, or are about to collect, a batch of data from the system:

$$Z^N = [y(1), u(1), y(2), u(2), \ldots, y(N), u(N)] \tag{7.6}$$

The problem we are faced with is to decide upon how to use the information contained in Z^N to select a proper value $\hat{\theta}_N$ of the parameter vector, and hence a proper member $\mathcal{M}(\hat{\theta}_N)$ in the set $\mathcal{M}^*$. Formally speaking, we have to determine a mapping from the data Z^N to the set $D_{\mathcal{M}}$:

$$Z^N \rightarrow \hat{\theta}_N \in D_{\mathcal{M}} \tag{7.7}$$

Such a mapping is a *parameter estimation method.*

Evaluating the Candidate Models

We are looking for a test by which the different models' ability to "describe" the observed data can be evaluated. We have stressed that the essence of a model is its prediction aspect, and we shall also judge its performance in this respect. Thus let the prediction error given by a certain model $\mathcal{M}(\theta_*)$ be given by

$$\varepsilon(t, \theta_*) = y(t) - \hat{y}(t|\theta_*) \tag{7.8}$$

When the data set Z^N is known, these errors can be computed for $t = 1, 2, \ldots, N$.

A "good" model, we say, is one that is good at predicting, that is, one that produces small prediction errors when applied to the observed data. Note that there is considerable flexibility in selecting various predictor functions, and this gives a

corresponding freedom in defining "good" models in terms of prediction performance. A guiding principle for parameter estimation thus is:

> Based on Z^t we can compute the prediction error $\varepsilon(t, \theta)$ using (7.8). At time $t = N$, select $\hat{\theta}_N$ so that the prediction errors $\varepsilon(t, \hat{\theta}_N), t = 1, 2, \ldots, N$, become as small as possible. (7.9)

The question is how to qualify what "small" should mean. In this chapter we shall describe two such approaches. One is to form a scalar-valued norm or criterion function that measures the size of ε. This approach is dealt with in Sections 7.2 to 7.4. Another approach is to demand that $\varepsilon(t, \hat{\theta}_N)$ be uncorrelated with a given data sequence. This corresponds to requiring that certain "projections" of $\varepsilon(t, \hat{\theta}_N)$ are zero and is further discussed in Sections 7.5 and 7.6.

7.2 MINIMIZING PREDICTION ERRORS

The prediction-error sequence in (7.8) can be seen as a vector in $\mathbf{R}^N$. The "size" of this vector could be measured using any norm in $\mathbf{R}^N$, quadratic or nonquadratic. This leaves a substantial amount of choices. We shall restrict the freedom somewhat by only considering the following way of evaluating "how large" the prediction-error sequence is: Let the prediction-error sequence be filtered through a stable linear filter $L(q)$:

$$\varepsilon_F(t, \theta) = L(q)\varepsilon(t, \theta), \qquad 1 \le t \le N \tag{7.10}$$

Then use the following norm:

$$V_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} \ell(\varepsilon_F(t, \theta)) \tag{7.11}$$

where $\ell(\cdot)$ is a scalar-valued (typically positive) function.

The function $V_N(\theta, Z^N)$ is, for given Z^N, a well-defined scalar-valued function of the model parameter θ. It is a natural measure of the validity of the model $\mathcal{M}(\theta)$. The estimate $\hat{\theta}_N$ is then defined by minimization of (7.11):

$$\hat{\theta}_N = \hat{\theta}_N(Z^N) = \arg\min_{\theta \in D_{\mathcal{M}}} V_N(\theta, Z^N) \tag{7.12}$$

Here arg min means "the minimizing argument of the function." If the minimum is not unique, we let arg min denote the set of minimizing arguments. The mapping (7.7) is thus defined implicitly by (7.12).

This way of estimating θ contains many well-known and much used procedures. We shall use the general term *prediction-error identification methods* (PEM) for the family of approaches that corresponds to (7.12). Particular methods, with specific "names" attached to themselves, are obtained as special cases of (7.12),

depending on the choice of $\ell(\cdot)$, the choice of prefilter $L(\cdot)$, the choice of model structure, and, in some cases, the choice of method by which the minimization is realized. We shall give particular attention to two especially well known members in the family (7.12) in the subsequent two sections. First, however, let us discuss some aspects on the choices of $L(q)$ and $\ell(\cdot)$ in (7.10) and (7.11). See also Section 15.2.

Choice of L

The effect of the filter L is to allow extra freedom in dealing with non-momentary properties of the prediction errors. Clearly, if the predictor is linear and time invariant, and y and u are scalars, then the result of filtering ε is the same as first filtering the input–output data and then applying the predictors.

The effect of L is best understood in a frequency-domain interpretation of (7.12), and a full discussion will be postponed to Chapter 13. It is clear, however, that by the use of L effects of high-frequency disturbances, not essential to the modeling problem, or slow drift terms and the like, can be removed. It also seems reasonable that certain properties of the models may be enhanced or suppressed by a properly selected L. L thus acts like *frequency weighting.*

The following particular aspect of the filtering (7.10) should be noted. If a model (7.3) is used, the filtered error $\varepsilon_F(t, \theta)$ is given by

$$\varepsilon_F(t, \theta) = L(q)\varepsilon(t, \theta) = [L^{-1}(q)H(q, \theta)]^{-1}[y(t) - G(q, \theta)u(t)] \qquad (7.13)$$

The effect of prefiltering is thus identical to changing the noise model from $H(q, \theta)$ *to*

$$\bar{H}_L(q, \theta) = L^{-1}(q)H(q, \theta) \qquad (7.14)$$

When we describe and analyze methods that employ general noise models in linear systems, we shall usually confine ourselves to $L(q) \equiv 1$, since the option of prefiltering is taken care of by the freedom in selecting $H(q, \theta)$. A discussion of the use and effects of $L(q)$ in practical terms will be given in Chapter 13.

Choice of ℓ

For the choice of $\ell(\cdot)$, a first candidate would be a quadratic norm:

$$\ell(\varepsilon) = \tfrac{1}{2}\varepsilon^2 \qquad (7.15)$$

and this is indeed a standard choice, which is convenient both for computation and analysis. Questions of robustness against bad data may, however, warrant other norms, which we shall discuss in some detail in Section 15.2. One may also conceive situations where the "best" norm is not known beforehand so that it is reasonable to parametrize the norm itself:

$$\ell(\varepsilon, \theta) \qquad (7.16)$$

Often the parametrization of the norm is then independent of the model parametrization:

$$\theta = \binom{\theta'}{\alpha}: \quad \ell(\varepsilon(t, \theta), \theta) = \ell(\varepsilon(t, \theta'), \alpha) \tag{7.17}$$

An exception to this case is given in Problem 7E.4.

Time-varying Norms

It may happen that measurements at different time instants are considered to be of varying reliability. The reason may be that the degree of noise corruption changes or that certain measurements are less representative for the system's properties. In such cases we are motivated to let the norm ℓ be time varying:

$$V_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^{N} \ell(\varepsilon(t, \theta), \theta, t) \tag{7.18}$$

In this way less reliable measurements can be associated with less weight in the criterion.

We shall frequently work with a criterion where the weighting is made explicitly by a weighting function $\beta(N, t)$:

$$V_N(\theta, Z^N) = \sum_{t=1}^{N} \beta(N, t)\ell(\varepsilon(t, \theta), \theta) \tag{7.19}$$

For fixed N, the N-dependence of $\beta(N, t)$ is of course immaterial. However, when estimates $\hat{\theta}_N$ for different N are compared, as for example in recursive identification (see Chapter 11), it becomes interesting to discuss how $\beta(N, t)$ varies with N. We shall return to this issue in Section 11.2.

Frequency-domain Interpretation of Quadratic Prediction-error Criteria for Linear Time-invariant Models

Let us consider the quadratic criterion error (7.12) and (7.15) for the standard linear model (7.3)

$$\begin{aligned} V_N(\theta, Z^N) &= \frac{1}{N} \sum_{t=1}^{N} \tfrac{1}{2}\varepsilon^2(t, \theta) \\ \varepsilon(t, \theta) &= H^{-1}(q, \theta)[y(t) - G(q, \theta)u(t)] \end{aligned} \tag{7.20}$$

Let $E_N(2\pi k/N, \theta)$, $k = 0, 1, \ldots, N-1$, be the DFT of $\varepsilon(t, \theta)$, $t = 1, 2, \ldots, N$:

$$E_N(2\pi k/N, \theta) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} \varepsilon(t, \theta)e^{-2\pi ikt/N}$$

Then, by Parseval's relation (2.44),

$$V_N(\theta, Z^N) = \frac{1}{N} \frac{1}{2} \sum_{k=0}^{N-1} |E_N(2\pi k/N, \theta)|^2 \tag{7.21}$$

Now let

$$w(t, \theta) = G(q, \theta)u(t)$$

Then the DFT of $w(t, \theta)$ is, according to Theorem 2.1,

$$W_N(\omega, \theta) = G(e^{i\omega}, \theta)U_N(\omega) + R_N(\omega)$$

with

$$|R_N(\omega)| \leq \frac{C}{\sqrt{N}}$$

The DFT of $s(t, \theta) = y(t) - w(t, \theta)$ then is

$$S_N(\omega, \theta) = Y_N(\omega) - G(e^{i\omega}, \theta)U_N(\omega) - R_N(\omega)$$

Finally,

$$\varepsilon(t, \theta) = H^{-1}(q, \theta)s(t, \theta)$$

has the DFT, again using Theorem 2.1,

$$E_N(\omega) = H^{-1}(e^{i\omega}, \theta)S_N(\omega, \theta) + \tilde{R}_N(\omega)$$

with

$$|\tilde{R}_N(\omega)| \leq \frac{C}{\sqrt{N}}$$

Inserting this into (7.21) gives

$$V_N(\theta, Z^N) = \frac{1}{N}\sum_{k=0}^{N-1}\frac{1}{2}|H(e^{2\pi ik/N}, \theta)|^{-2}|Y_N(2\pi k/N) - G(e^{2\pi ik/N}, \theta)U_N(2\pi k/N)|^2 + \bar{R}_N$$

with $|\bar{R}_N| \leq C/\sqrt{N}$ or, using the definition of the ETFE $\hat{\hat{G}}_N$ in (6.24),

$$V_N(\theta, Z^N) = \frac{1}{N}\sum_{k=0}^{N-1}\left\{\frac{1}{2}|\hat{\hat{G}}_N(e^{2\pi ik/N}) - G(e^{2\pi ik/N}, \theta)|^2 Q_N(2\pi k/N, \theta) + \bar{R}_N\right\} \tag{7.22}$$

with

$$Q_N(\omega, \theta) = \frac{|U_N(\omega)|^2}{|H(e^{i\omega}, \theta)|^2} \tag{7.23}$$

First notice that, apart from the remainder term $\bar{R}_N$, the expression (7.22) coincides with the weighted least-squares criterion for a model:

$$\hat{\hat{G}}_N(e^{2\pi ik/N}) = G(e^{2\pi ik/N}, \theta) + v(k) \tag{7.24}$$

Compare with (II.96) and (II.97). According to Lemma 6.1, the variance of $v(k)$ is, asymptotically, $\Phi_v(2\pi k/N)/|U_N(2\pi k/N)|^2$, so the weighting coefficient $Q_N(\omega, \theta)$ is the inverse variance, which is optimal for linear regressions, according to (II.65). In (7.23) the unknown noise spectrum $\Phi_v(\omega)$ is replaced by the model noise spectrum $|H(e^{i\omega}, \theta)|^2$. Consequently, the prediction-error methods can be seen as methods of fitting the ETFE to the model transfer function with a weighted norm,

corresponding to the model signal-to-noise ratio at the frequency in question. For notational reasons, it is instructive to rewrite the sum (7.22) approximately as an integral:

$$V_N(\theta, Z^N) \approx \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{2}|\hat{G}_N(e^{i\omega}) - G(e^{i\omega}, \theta)|^2 Q_N(\omega, \theta)\, d\omega \qquad (7.25)$$

The shift of integration interval from $(0, 2\pi)$ to $(-\pi, \pi)$ is possible since the integrand is periodic.

With this interpretation we have described the prediction-error estimate as an alternative way of smoothing the ETFE, showing a strong conceptual relationship to the spectral analysis methods of Section 6.4. See Problem 7G.2 for a direct tie.

When we specialize to the case of a time series [no input and $G(q, \theta) \equiv 0$], the criterion (7.25) takes the form

$$V_N(\theta, Z^N) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|\frac{Y_N(\omega)}{H(e^{i\omega}, \theta)}\right|^2 d\omega \qquad (7.26)$$

Such parametric estimators of spectra are known as "Whittle-type estimators," after Whittle (1951).

Multivariable Systems (*)

For multioutput systems, the counterpart of the quadratic criterion is

$$\ell(\varepsilon) = \tfrac{1}{2}\varepsilon^T \Lambda^{-1} \varepsilon \qquad (7.27)$$

for some symmetric, positive semidefinite $p \times p$ matrix Λ that weighs together the relative importance of the components of ε.

One might discuss what is the best choice of norm Λ. We shall do that in some detail in Section 15.2. Here we only remark that, just as in (7.16), the parameter vector θ could be extended to include components of Λ, and the function ℓ will then be an appropriate function of θ.

As a variant of the criterion (7.11), where a scalar $\ell(\varepsilon)$ is formed for each t, we could first form the $p \times p$ matrix

$$Q_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} \varepsilon(t, \theta)\varepsilon^T(t, \theta) \qquad (7.28)$$

and let the criterion be a scalar-valued function of this matrix:

$$V_N(\theta, Z^N) = h(Q_N(\theta, Z^N)) \qquad (7.29)$$

The criterion (7.27) is then obtained by

$$h(Q) = \tfrac{1}{2}\operatorname{tr}(Q\Lambda^{-1}) \qquad (7.30)$$

7.3 LINEAR REGRESSIONS AND THE LEAST-SQUARES METHOD

Linear Regressions

We found in both Sections 4.2 and 5.2 that linear regression model structures are very useful in describing basic linear and nonlinear systems. The linear regression employs a predictor (5.34)

$$\hat{y}(t|\theta) = \varphi^T(t)\theta + \mu(t) \tag{7.31}$$

that is linear in θ. Here φ is the vector of regressors, the *regression vector.* Recall that for the ARX structure (4.7) we have

$$\varphi(t) = [-y(t-1)\ -y(t-2) \dots -y(t-n_a)\ u(t-1) \dots u(t-n_b)]^T \tag{7.32}$$

In (7.31), $\mu(t)$ is a known data-dependent vector. For notational simplicity we shall take $\mu(t) = 0$ in the remainder of this section; it is quite straightforward to include it. See Problem 7D.1.

Linear regression forms a standard topic in statistics. The reader could consult Appendix II for a refresher of basic properties. The present section can, however, be read independently of Appendix II.

Least-squares Criterion

With (7.31) the prediction error becomes

$$\varepsilon(t, \theta) = y(t) - \varphi^T(t)\theta$$

and the criterion function resulting from (7.10) and (7.11), with $L(q) = 1$ and $\ell(\varepsilon) = \frac{1}{2}\varepsilon^2$, is

$$V_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} \tfrac{1}{2}[y(t) - \varphi^T(t)\theta]^2 \tag{7.33}$$

This is the *least-squares criterion* for the linear regression (7.31). The unique feature of this criterion, developed from the linear parametrization and the quadratic criterion, is that it is a quadratic function in θ. Therefore, it can be minimized analytically, which gives, provided the indicated inverse exists,

$$\hat{\theta}_N^{\text{LS}} = \arg\min V_N(\theta, Z^N) = \left[\frac{1}{N}\sum_{t=1}^{N} \varphi(t)\varphi^T(t)\right]^{-1} \frac{1}{N}\sum_{t=1}^{N} \varphi(t)y(t) \tag{7.34}$$

the *least-squares estimate (LSE)* (see Problem 7D.2).

Introduce the $d \times d$ matrix

$$R(N) = \frac{1}{N}\sum_{t=1}^{N} \varphi(t)\varphi^T(t) \tag{7.35}$$

and the d-dimensional column vector

$$f(N) = \frac{1}{N}\sum_{t=1}^{N} \varphi(t)y(t) \tag{7.36}$$

In the case (7.32), $\varphi(t)$ contains lagged input and output variables, and the entries of the quantities (7.35) and (7.36) will be of the form

$$[R(N)]_{ij} = \frac{1}{N}\sum_{t=1}^{N} y(t-i)y(t-j), \qquad 1 \le i, j \le n_a$$

and similar sums of $u(t-r)\cdot u(t-s)$ or $u(t-r)\cdot y(t-s)$ for the other entries of $R(N)$. That is, they will consist of estimates of the covariance functions of $\{y(t)\}$ and $\{u(t)\}$. The LSE can thus be computed using only such estimates and is therefore related to correlation analysis, as described in Section 6.1.

Properties of the LSE

The least-squares method is a special case of the prediction-error identification method (7.12). An analysis of its properties is therefore contained in the general treatment in Chapters 8 and 9. It is, however, useful to include a heuristic investigation of the LSE at this point.

Suppose that the observed data actually have been generated by

$$y(t) = \varphi^T(t)\theta_0 + v_0(t) \tag{7.37}$$

for some sequence $\{v_0(t)\}$. We may think of θ_0 as a "true value" of the parameter vector. Inserting (7.37) into (7.34) gives

$$\begin{aligned}\hat{\theta}_N^{LS} &= [R(N)]^{-1}\frac{1}{N}\sum_{t=1}^{N}\varphi(t)[\varphi^T(t)\theta_0 + v_0(t)] \\ &= \theta_0 + [R(N)]^{-1}\frac{1}{N}\sum_{t=1}^{N}\varphi(t)v_0(t)\end{aligned} \tag{7.38}$$

Desired properties of $\hat{\theta}_N^{LS}$ would be that

1. It is close to θ_0.
2. It converges to θ_0 as N tends to infinity.

We first note that if $v_0(t)$ in (7.37) is small compared to $\varphi(t)$ then the error term

$$[R(N)]^{-1}\frac{1}{N}\sum_{t=1}^{N}\varphi(t)v_0(t)$$

will be small, and thus $\hat{\theta}_N^{LS}$ will be close to θ_0. To investigate what happens when N tends to infinity, it is convenient to assume that $\{v_0(t)\}$ is a realization of a stationary stochastic process and to specialize to (7.32). Assume that the input is quasi-stationary so that sums like

$$\hat{R}_u^N(\tau) = \frac{1}{N}\sum_{t=1}^{N} u(t)u(t-\tau) \to R_u(\tau) = \bar{E}u(t-\tau)u(t)$$

converge when N tends to infinity. Then the matrix $R(N)$ (which is made up from such sums and related ones) will converge (with probability 1):

$$R(N) \to R^*, \qquad (= \bar{E}\varphi(t)\varphi^T(t)) \text{ as } N \to \infty$$

Also, under weak conditions with probability 1,

$$\frac{1}{N}\sum_{t=1}^{N} \varphi(t)v_0(t) \to h^*, \qquad (= \bar{E}\varphi(t)v_0(t)) \text{ as } N \to \infty$$

(Compare with Theorem 2.3.) Thus

$$\hat{\theta}_N \to \theta_0 + (R^*)^{-1}h^*, \qquad \text{as } N \to \infty \tag{7.39}$$

provided R^* is nonsingular. For the LSE to be *consistent,* that is, for $\hat{\theta}_N$ to converge to θ_0, we thus have to require:

(i) R^* is nonsingular. This will typically be the case, for example, if $\{u(t)\}$ and $\{v_0(t)\}$ are independent and the $m \times m$ matrix, whose i, j entry is $R_u(i-j)$, is nonsingular. In this case the input is said to be *persistently exciting* of order n_b (a thorough discussion of this concept is given in Section 14.2).

(ii) $h^* = 0$. This will be the case if either

(iia) $\{v_0(t)\}$ is a sequence of independent random variables with zero mean values (white noise). Then $v_0(t)$ does not depend on what has happened up to time $t-1$, and hence $E\varphi(t)v_0(t) = 0$.

or

(iib) The input sequence $\{u(t)\}$ is independent of the zero-mean noise sequence $\{v_0(t)\}$ and $n_a = 0$ in (7.32). Then $\varphi(t)$ contains only u terms and hence $E\varphi(t)v_0(t) = 0$.

When $n_a > 0$ so that $\varphi(t)$ contains $y(k)$, $t - n_a \le k \le t-1$, and $\{v_0(t)\}$ is not white noise, then (usually) $Ev_0(t)\varphi(t) \neq 0$. This follows since $\varphi(t)$ contains $y(t-1)$, while $y(t-1)$ contains the term $v_0(t-1)$ that is correlated with $v_0(t)$. Therefore, we may expect consistency usually only in cases (iia) and (iib).

In the cases (i) and (iia) it can also be shown (see Chapter 9) that the random variable

$$\sqrt{N}(\hat{\theta}_N - \theta_0)$$

converges in distribution to the normal distribution with zero mean and covariance matrix $\lambda_0[R^*]^{-1}$, where λ_0 is the variance of $v_0(t)$. *Experiment design* issues [e.g., the choice of the properties of $\{u(t)\}$] therefore deal with the problem of making R^* "large" subject to given constraints. Such issues are discussed in Chapter 14.

Weighted Least Squares

Just as in (7.18) and (7.19), the different measurements could be assigned different weights in the least-squares criterion:

$$V_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} \alpha_t [y(t) - \varphi^T(t)\theta]^2 \tag{7.40}$$

or

$$V_N(\theta, Z^N) = \sum_{t=1}^{N} \beta(N, t)[y(t) - \varphi^T(t)\theta]^2 \tag{7.41}$$

The expression for the resulting estimate is quite analogous to (7.34):

$$\hat{\theta}_N^{\text{LS}} = \left[\sum_{t=1}^{N} \beta(N, t)\varphi(t)\varphi^T(t)\right]^{-1} \sum_{t=1}^{N} \beta(N, t)\varphi(t)y(t) \tag{7.42}$$

Multivariable Case (*)

If the output $y(t)$ is a p-vector and the norm (7.27) is used, the LS criterion takes the form

$$V_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} \tfrac{1}{2}[y(t) - \varphi^T(t)\theta]^T \Lambda^{-1}[y(t) - \varphi^T(t)\theta] \tag{7.43}$$

This gives the estimate

$$\hat{\theta}_N^{\text{LS}} = \left[\frac{1}{N}\sum_{t=1}^{N} \varphi(t)\Lambda^{-1}\varphi^T(t)\right]^{-1} \frac{1}{N}\sum_{t=1}^{N} \varphi(t)\Lambda^{-1}y(t) \tag{7.44}$$

In case we use the particular parametrization (4.53) with $\boldsymbol{\theta}$ as an $r \times p$ matrix,

$$\hat{y}(t|\theta) = \boldsymbol{\theta}^T\varphi(t) \tag{7.45}$$

the LS criterion becomes

$$V_N(\boldsymbol{\theta}, Z^N) = \frac{1}{N}\sum_{t=1}^{N} \|y(t) - \boldsymbol{\theta}^T\varphi(t)\|^2 \tag{7.46}$$

with the estimate

$$\hat{\boldsymbol{\theta}}_N^{\text{LS}} = \left[\frac{1}{N}\sum_{t=1}^{N} \varphi(t)\varphi^T(t)\right]^{-1} \frac{1}{N}\sum_{t=1}^{N} \varphi(t)y^T(t) \tag{7.47}$$

(see problem 7D.2). The expression (7.47) brings out the advantages of the structure (7.45): To determine the $r \times p$ estimate $\hat{\boldsymbol{\theta}}_N$, it is sufficient to invert an $r \times r$ matrix. In (7.44) θ is a $p \cdot r$ vector and the matrix inversion involves a $pr \times pr$ matrix.

Colored Equation-error Noise (*)

The LS method has many advantages, the most important one being that the global minimum of (7.33) can be found efficiently and unambiguously (no local minima other than global ones exist). Its main shortcoming relates to the asymptotic properties quoted previously: If, in a difference equation,

$$y(t) + a_1 y(t-1) + \cdots + a_{n_a} y(t-n_a) = b_1 u(t-1) + \cdots + b_{n_b} u(t-n_b) + v(t) \qquad (7.48)$$

the equation error $v(t)$ is not white noise, then the LSE will not converge to the true values of a_i and b_i. To deal with this problem, we may incorporate further modeling of the equation error $v(t)$ as discussed in Section 4.2, let us say

$$v(t) = \kappa(q)e(t) \qquad (7.49)$$

with e white and κ a linear filter. Models employing (7.49) will typically take us out from the LS environment, except in two cases, which we now discuss.

Known noise properties: If in (7.48) and (7.49) a_i and b_i are unknown, but κ is a known filter (not too realistic a situation), we have

$$A(q)y(t) = B(q)u(t) + \kappa(q)e(t) \qquad (7.50)$$

Filtering (7.50) through the filter $\kappa^{-1}(q)$ gives

$$A(q)y_F(t) = B(q)u_F(t) + e(t) \qquad (7.51)$$

where

$$y_F(t) = \kappa^{-1}(q)y(t), \qquad u_F(t) = \kappa^{-1}(q)u(t) \qquad (7.52)$$

Since e is white, the LS method can be applied to (7.51) without problems. Notice that this is equivalent to applying the filter $L(q) = \kappa^{-1}(q)$ in (7.10).

High-order models: Suppose that the noise v can be well described by $\kappa(q) = 1/D(q)$ in (7.49), where $D(q)$ is a polynomial of degree r. [That is, $v(t)$ is supposed to be an autoregressive (AR) process of order r.] This gives

$$A(q)y(t) = B(q)u(t) + \frac{1}{D(q)}e(t) \qquad (7.53)$$

or

$$A(q)D(q)y(t) = B(q)D(q)u(t) + e(t) \qquad (7.54)$$

Applying the LS method to (7.54) with orders $n_A = n_a + r$ and $n_B = n_b + r$ gives, since e is white, consistent estimates of AD and BD. Hence the transfer function from u to y,

$$\frac{B(q)D(q)}{A(q)D(q)} = \frac{B(q)}{A(q)}$$

is correctly estimated. This approach was called *repeated least squares* in Åström and Eykhoff (1971). See also Söderström (1975b) and Stoica (1976).

7.4 A STATISTICAL FRAMEWORK FOR PARAMETER ESTIMATION AND THE MAXIMUM LIKELIHOOD METHOD

So far we have not appealed to any statistical arguments for the estimation of θ. In fact, our framework of fitting models to data makes sense regardless of a stochastic setting of the data. It is, however, useful and instructive at this point to briefly describe basic aspects of statistical parameter estimation and relate them to our framework.

Estimators and the Principle of Maximum Likelihood

The area of statistical inference, as well as that of system identification and parameter estimation, deals with the problem of extracting information from observations that themselves could be unreliable. The observations are then described as realizations of stochastic variables. Suppose that the observations are represented by the random variable $y^N = (y(1), y(2), \ldots, y(N))$ that takes values in $\mathbf{R}^N$. The probability density function (PDF) of y^N is supposed to be

$$f(\theta; x_1, x_2, \ldots, x_N) = f_y(\theta; x^N) \tag{7.55}$$

That is,

$$P(y^N \in A) = \int_{x^N \in A} f_y(\theta; x^N) dx^N \tag{7.56}$$

In (7.55), θ is a d-dimensional parameter vector that describes properties of the observed variable. These are supposed to be unknown, and the purpose of the observation is in fact to estimate the vector θ using y^N. This is accomplished by an *estimator,*

$$\hat{\theta}(y^N) \tag{7.57}$$

which is a function from $\mathbf{R}^N$ to $\mathbf{R}^d$. If the observed value of y^N is y_*^N, then consequently the resulting estimate is $\hat{\theta}_* = \hat{\theta}(y_*^N)$.

Many such estimator functions are possible. A particular one that maximizes the probability of the observed event is the celebrated maximum likelihood estimator, introduced by Fisher (1912). It can be defined as follows: The joint probability density function for the random vector to be observed is given by (7.55). The probability that the realization (= observation) indeed should take the value y_*^N is thus proportional to

$$f_y(\theta; y_*^N)$$

This is a deterministic function of θ once the numerical value y_*^N is inserted. This function is called the *likelihood function.* It reflects the "likelihood" that the observed event should indeed take place. A reasonable estimator of θ could then be to select it so that the observed event becomes "as likely as possible." That is, we seek

$$\hat{\theta}_{\mathrm{ML}}(y_*^N) = \arg\max_{\theta} f_y(\theta; y_*^N) \tag{7.58}$$

where the maximization is performed for fixed y_*^N. This function is known as the *maximum likelihood estimator* (MLE).

An Example

Let $y(i), i = 1, \ldots, N$, be independent random variables with normal distribution with (unknown) means θ_0 (independent of i) and (known) variances λ_i:

$$y(i) \in N(\theta_0, \lambda_i) \tag{7.59}$$

A common estimator of θ_0 is the sample mean:

$$\hat{\theta}_{\text{SM}}(y^N) = \frac{1}{N}\sum_{i=1}^{N} y(i) \tag{7.60}$$

To calculate the MLE, we start by determining the joint PDF (7.55) for the observations. Since the PDF for $y(i)$ is

$$\frac{1}{\sqrt{2\pi\lambda_i}} \exp\left[-\frac{(x_i - \theta)^2}{2\lambda_i}\right]$$

and the $y(i)$ are independent, we have

$$f_y(\theta; x^N) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\lambda_i}} \exp\left[-\frac{(x_i - \theta)^2}{2\lambda_i}\right] \tag{7.61}$$

The likelihood function is thus given by $f_y(\theta; y^N)$. Maximizing the likelihood function is the same as maximizing its logarithm. Thus

$$\begin{aligned}\hat{\theta}_{\text{ML}}(y^N) &= \arg\max_{\theta} \log f_y(\theta; y^N) \\ &= \arg\max_{\theta} \left\{-\frac{N}{2}\log 2\pi - \sum_{i=1}^{N}\frac{1}{2}\log\lambda_i - \frac{1}{2}\sum_{i=1}^{N}\frac{(y(i) - \theta)^2}{\lambda_i}\right\}\end{aligned} \tag{7.62}$$

from which we find

$$\hat{\theta}_{\text{ML}}(y^N) = \frac{1}{\sum_{i=1}^{N}(1/\lambda_i)} \sum_{i=1}^{N} \frac{y(i)}{\lambda_i} \tag{7.63}$$

Relationship to the Maximum A Posteriori (MAP) Estimate

The *Bayesian approach* gives a related but conceptually different treatment of the parameter estimation problem. In the Bayesian approach the parameter itself is thought of as a random variable. Based on observations of other random variables that are correlated with the parameter, we may infer information about its value. Suppose that the properties of the observations can be described in terms of a parameter vector θ. With a Bayesian view we thus consider θ to be a random vector

with a certain prior distribution ("prior" means before the observations have been made). The observations y^N are obviously correlated with this θ. After the observations have been obtained, we then ask for the posterior PDF for θ. From this posterior PDF, different estimates of θ can be determined, for example, the value for which the PDF attains its maximum ("the most likely value"). This is known as the *maximum a posteriori (MAP) estimate.*

Suppose that the conditional PDF for y^N, given θ, is

$$f_y(\theta; x^N) = P(y^N = x^N|\theta)$$

and that the prior PDF for θ is

$$g_\theta(z) = P(\theta = z)$$

[Here $P(A|B)$ = the conditional probability of the event A given B. We also allowed somewhat informal notation.] Using Bayes's rule (I.10) and with some abuse of notation, we thus find the posterior PDF for θ, i.e., the conditional PDF for θ, given the observations:

$$P(\theta|y^N) = \frac{P(y^N|\theta) \cdot P(\theta)}{P(y^N)} \sim f_y(\theta; y^N) \cdot g_\theta(\theta) \tag{7.64}$$

The posterior PDF as a function of θ is thus proportional to the likelihood function multiplied by the prior PDF. Often the prior PDF has an insignificant influence. Then the MAP estimate

$$\hat{\theta}_{\text{MAP}}(y^N) = \arg \max_\theta \{f_y(\theta; y^N) \cdot g_\theta(\theta)\} \tag{7.65}$$

is close to the MLE (7.58).

Cramér–Rao Inequality

The quality of an estimator can be assessed by its mean-square error matrix:

$$P = E[\hat{\theta}(y^N) - \theta_0][\hat{\theta}(y^N) - \theta_0]^T \tag{7.66}$$

Here θ_0 denotes the "true value" of θ, and (7.66) is evaluated under the assumption that the PDF of y^N is $f_y(\theta_0; y^N)$.

We may be interested in selecting estimators that make P small. It is then interesting to note that there is a lower limit to the values of P that can be obtained with various unbiased estimators. This is the so called *Cramér–Rao inequality:*

> Let $\hat{\theta}(y^N)$ be an estimator of θ such that $E\,\hat{\theta}(y^N) = \theta_0$, where E evaluates the mean, assuming that the PDF of y^N is $f_y(\theta_0; y^N)$ (to hold for all values of θ_0), and suppose that y^N may take values in a subset of $\mathbf{R}^N$, whose boundary does not depend on θ. Then
>
> $$E[\hat{\theta}(y^N) - \theta_0][\hat{\theta}(y^N) - \theta_0]^T \geq M^{-1} \tag{7.67}$$

where

$$M = E\left[\frac{d}{d\theta}\log f_y(\theta; y^N)\right]\left[\frac{d}{d\theta}\log f_y(\theta; y^N)\right]^T\Big|_{\theta=\theta_0}$$
$$= -E\frac{d^2}{d\theta^2}\log f_y(\theta; y^N)\Big|_{\theta=\theta_0} \tag{7.68}$$

Since θ is a d-dimensional vector, $(d/d\theta)\log f_y(\theta; y^N)$ is a d-dimensional column vector and the Hessian $(d^2/d\theta^2)\log f_y(\theta; y^N)$ is a $d \times d$ matrix. This matrix M is known as the *Fisher information matrix*. Notice that the evaluation of M normally requires knowledge of θ_0, so the exact value of M may not be available to the user.

A proof of the Cramér–Rao inequality is given in Appendix 7A.

Asymptotic Properties of the MLE

It is often difficult to exactly calculate properties of an estimator, such as (7.66). Therefore, limiting properties as the sample size (in this case the number N) tends to infinity are calculated instead. Classical such results for the MLE in case of independent observations were obtained by Wald (1949) and Cramér (1946):

> Suppose that the random variables $\{y(i)\}$ are independent and identically distributed, so that
>
> $$f_y(\theta; x_1, \ldots, x_N) = \prod_{i=1}^{N} f_{y(i)}(\theta, x_i)$$
>
> Suppose also that the distribution of y^N is given by $f_y(\theta_0; x^N)$ for some value θ_0. Then the random variable $\hat{\theta}_{ML}(y^N)$ tends to θ_0 with probability 1 as N tends to infinity, and the random variable
>
> $$\sqrt{N}\,[\hat{\theta}_{ML}(y^N) - \theta_0]$$
>
> converges in distribution to the normal distribution with zero mean and covariance matrix given by the Cramér–Rao lower bound [M^{-1} in (7.67) and (7.68)].

In Chapters 8 and 9 we will establish that these results also hold when the ML estimator is applied to dynamical systems. In this sense the MLE is thus the best possible estimator. Let it, however, also be said that the MLE sometimes has been criticized for less good small sample properties and that there are other ways to assess the quality of an estimator than (7.66).

Probabilistic Models of Dynamical Systems

Suppose that the models in the model structure we have chosen in Section 7.1 include both a predictor function and an assumed PDF for the associated prediction errors, as described in Section 5.4:

$$\mathcal{M}(\theta): \quad \hat{y}(t|\theta) = g(t, Z^{t-1}; \theta)$$
$$\varepsilon(t, \theta) = y(t) - \hat{y}(t|\theta) \text{ are independent and have the PDF } f_e(x, t; \theta) \tag{7.69}$$

Recall that we term a model like (7.69) that includes a PDF for ε a (complete) *probabilistic model.*

Likelihood Function for Probabilistic Models of Dynamical Systems

We note that, according to the model (7.69), the output is generated by

$$y(t) = g(t, Z^{t-1}; \theta) + \varepsilon(t, \theta) \tag{7.70}$$

where $\varepsilon(t, \theta)$ has the PDF $f_e(x, t; \theta)$. The joint PDF for the observations y^N (given the deterministic sequence u^N) is then given by Lemma 5.1. By replacing the dummy variables x_i by the corresponding observations $y(i)$, we obtain the likelihood function:

$$\begin{aligned} \bar{f}_y(\theta; y^N) &= \prod_{t=1}^{N} f_e(y(t) - g(t, Z^{t-1}; \theta), t; \theta) \\ &= \prod_{t=1}^{N} f_e(\varepsilon(t, \theta), t; \theta) \end{aligned} \tag{7.71}$$

Maximizing this function is the same as maximizing

$$\frac{1}{N} \log \bar{f}_y(\theta; y^N) = \frac{1}{N} \sum_{t=1}^{N} \log f_e(\varepsilon(t,\theta), t; \theta) \tag{7.72}$$

If we define

$$\ell(\varepsilon, \theta, t) = -\log f_e(\varepsilon, t; \theta) \tag{7.73}$$

we may write

$$\hat{\theta}_{\mathrm{ML}}(y^N) = \arg\min_{\theta} \frac{1}{N} \sum_{t=1}^{N} \ell(\varepsilon(t,\theta), \theta, t) \tag{7.74}$$

The maximum likelihood method can thus be seen as a special case of the prediction-error criterion (7.12).

It is worth stressing that (7.73) and (7.74) give the exact maximum likelihood method for the posed problem. It is sometimes pointed out that the exact likelihood function is quite complicated for time-series problems and that one has to resort to approximations of it (e.g., Kashyap and Rao, 1976; Akaike, 1973; Dzhaparidze and Yaglom, 1983). This is true in certain cases. The reason is that it may be difficult to put, say, an ARMA model in the predictor form (7.69) (it will typically require time-varying Kalman predictors). The problem is therefore related to finding the exact predictor and is not a problem with the ML method as such. When we employ

time-invariant predictors, we implicitly assume all previous observations to be known [see (3.24)] and typically replace the corresponding initial values by zero. Then it is appropriate to interpret the likelihood function as *conditional* w.r.t. these values and to call the method a *conditional ML method* (e.g., Kashyap and Rao, 1976).

Gaussian Special Case

When the prediction errors are assumed to be Gaussian with zero mean values and (t-independent) covariances λ, we have

$$\ell(\varepsilon,\theta,t) = -\log f_e(\varepsilon,t;\theta) = \text{const} + \frac{1}{2}\log\lambda + \frac{1}{2}\frac{\varepsilon^2}{\lambda} \tag{7.75}$$

If λ is known, then (7.75) is equivalent to the quadratic criterion (7.15). If λ is unknown, (7.75) is an example of a parameterized norm criterion (7.16). Depending on the underlying model structure, λ may or may not be parametrized independently of the predictor parameters. See Problem 7E.4 for an illustration of this. Compare also Problem 7E.7.

Fisher Information Matrix and the Cramér–Rao Bound for Dynamical Systems

Having established the log likelihood function in (7.72) for a model structure, we can compute the information matrix (7.68). For simplicity, we then assume that the PDF f_e is known (θ independent) and time invariant. Let $\ell_0(\varepsilon) = -\log f_e(\varepsilon)$. Hence

$$\frac{d}{d\theta}\log\bar{f}_y(\theta;y^N) = \sum_{t=1}^{N}\ell_0'(\varepsilon(t,\theta))\cdot\psi(t,\theta)$$

where, as in (4.118c),

$$\psi(t,\theta) = \frac{d}{d\theta}\hat{y}(t|\theta) = -\frac{d}{d\theta}\varepsilon(t,\theta), \qquad \text{[a } d\text{-dimensional column vector]}$$

Also, ℓ_0' is the derivative of $\ell_0(\varepsilon)$ w.r.t. ε. To find the Fisher information matrix, we now evaluate the expectation of

$$\frac{d}{d\theta}\log\bar{f}_y(\theta;y^N)\left[\frac{d}{d\theta}\log\bar{f}_y(\theta;y^N)\right]^T$$

at θ_0 under the assumption that the true PDF for y^N indeed is $\bar{f}_y(\theta_0;y^N)$. The latter statement means that $\varepsilon(t,\theta_0) = e_0(t)$ will be treated as a sequence of independent random variables with PDF's $f_e(x)$. Call this expectation M_N. Thus

$$M_N = E\sum_{t=1}^{N}\sum_{s=1}^{N} \ell_0'(e_0(t))\ell_0'(e_0(s))\psi(t,\theta_0)\psi^T(s,\theta_0)$$
$$= \sum_{t=1}^{N} E[\ell_0'(e_0(t))]^2 \cdot E\psi(t,\theta_0)\psi^T(t,\theta_0)$$

since $e_0(t)$ and $e_0(s)$ are independent for $s \neq t$. We also have $\ell_0'(x) = [\log f_e(x)]' = f_e'(x)/f_e(x)$, and

$$E[\ell_0'(e_0(t))]^2 = \int \frac{[f_e'(x)]^2}{f_e^2(x)} \cdot f_e(x)dx$$
$$= \int_{-\infty}^{\infty} \frac{[f_e'(x)]^2}{f_e(x)}dx \triangleq \frac{1}{\kappa_0} \tag{7.76}$$

If $e_0(t)$ is Gaussian with variance λ_0, it is easy to verify that $\kappa_0 = \lambda_0$. Hence

$$M_N = \frac{1}{\kappa_0} \cdot \sum_{t=1}^{N} E\psi(t,\theta_0)\psi^T(t,\theta_0) \tag{7.77}$$

Now the Cramér–Rao inequality tells us that for *any unbiased estimator* $\hat{\theta}_N$ of θ (i.e., estimators such that $E\hat{\theta}_N = \theta_0$ regardless of the true value θ_0) we must have

$$\text{Cov}\,\hat{\theta}_N \geq M_N^{-1} \tag{7.78}$$

Notice that this bound applies for any N and for all parameter estimation methods. We thus have

$$\text{Cov}\,\hat{\theta}_N \geq \kappa_0 \left[\sum_{t=1}^{N} E\psi(t,\theta_0)\psi^T(t,\theta_0)\right]^{-1} \tag{7.79}$$
$$\kappa_0 = \lambda_0 \text{ for Gaussian innovations}$$

Multivariable Gaussian Case (*)

When the prediction errors are p-dimensional and jointly Gaussian with zero mean and covariance matrices Λ, we obtain from the multivariable Gaussian distribution

$$\ell(\varepsilon, t; \theta) = \text{const} + \frac{1}{2}\log\det\Lambda + \frac{1}{2}\varepsilon^T\Lambda^{-1}\varepsilon \tag{7.80}$$

Then the negative logarithm of the likelihood function takes the form

$$V_N(\theta, \Lambda, Z^N) = \text{const} + \frac{N}{2}\log\det\Lambda + \frac{1}{2}\sum_{t=1}^{N} \varepsilon^T(t,\theta)\Lambda^{-1}\varepsilon(t,\theta) \tag{7.81}$$

If the $p \times p$ covariance matrix Λ is fully unknown and not parametrized through θ, it is possible to minimize (7.81) analytically with respect to Λ for every fixed θ:

$$\arg\min_{\Lambda} V_N(\theta, \Lambda, Z^N) = \hat{\Lambda}_N(\theta) = \frac{1}{N}\sum_{t=1}^{N} \varepsilon(t, \theta)\varepsilon^T(t, \theta) \tag{7.82}$$

Then

$$\begin{aligned}\hat{\theta}_N &= \arg\min_{\theta} V_N(\theta, \hat{\Lambda}_N(\theta), Z^N) \\ &= \arg\min_{\theta} \left[\tfrac{1}{2}\log\det\hat{\Lambda}_N(\theta) + \tfrac{1}{2}p\right]\end{aligned} \tag{7.83}$$

(see problem 7D.3) where $p = \dim \varepsilon$. Hence we may in this particular case use the criterion

$$\hat{\theta}_N = \arg\min_{\theta} \det\left[\frac{1}{N}\sum_{t=1}^{N} \varepsilon(t, \theta)\varepsilon^T(t, \theta)\right] \tag{7.84}$$

With this we have actually been led to a criterion of the type (7.29) to (7.30) with $h(A) = \det A$.

Information and Entropy Measures (*)

In (5.36) and (5.37) we gave a general formulation of a model as an assumed PDF for the observations Z^t:

$$\bar{f}_m(t, Z^t) \tag{7.85}$$

Let $\bar{f}_0(t, Z^t)$ denote the true PDF for the observations. The agreement between two PDF's can be measured in terms of the *Kullback–Leibler information distance* (Kullback and Leibler, 1951):

$$I(\bar{f}_0; \bar{f}_m) = \int \bar{f}_0(t, x^t) \log \frac{\bar{f}_0(t, x^t)}{\bar{f}_m(t, x^t)}\, dx^t \tag{7.86}$$

Here we use x^t as an integration variable for Z^t. This distance is also the *negative entropy* of $\bar{f}_0$ with respect to $\bar{f}_m$:

$$S(\bar{f}_0; \bar{f}_m) = -I(\bar{f}_0; \bar{f}_m) \tag{7.87}$$

An attractive formulation of the identification problem is to *look for a model that maximizes the entropy with respect to the true system* or, alternatively, *minimizes the information distance to the true system.* This formulation has been pursued by Akaike in a number of interesting contributions Akaike (1972, 1974a, 1981).

With a parametrized set of models $\bar{f}_{\mathcal{M}(\theta)}(t, Z^t) = \bar{f}(\theta; t, Z^t)$, we would thus solve

$$\hat{\theta}_N = \arg\min_{\theta} I(\bar{f}_0(N, Z^N); \bar{f}(\theta; N, Z^N)) \tag{7.88}$$

The information measure can be written

$$I(\bar{f}_0; \bar{f}) = -\int \log[\bar{f}(\theta; N, x^N)] \cdot \bar{f}_0(N, x^N) dx^N$$

$$+ \int \log[\bar{f}_0(N, x^N)] \cdot \bar{f}_0(N, x^N) dx^N$$

$$= -E_0 \log \bar{f}(\theta; N, Z^N) + \theta\text{-independent terms}$$

where E_0 denotes expectation with respect to the true system.

The problem (7.88) is thus the same as

$$\hat{\theta}_N = \arg \min_{\theta} [-E_0 \log \bar{f}(\theta; N, Z^N)] \tag{7.89}$$

The problem here is of course that the expectation is not computable since the true PDF is unknown. A simple estimate of the expectation is to replace it by the observation

$$E_0 \log \bar{f}(\theta; N, Z^N) \approx \log \bar{f}(\theta; N, Z^N) \tag{7.90}$$

This gives the log likelihood function for the problem and (7.89) then equals the MLE. The ML approach to identification can consequently also be interpreted as a maximum entropy strategy or a minimum information distance method.

The distance between the resulting model and the true system thus is

$$I(\bar{f}_0(N, Z^N); \bar{f}(\hat{\theta}_N; N, Z^N)) \tag{7.91}$$

This is a random variable, since $\hat{\theta}_N$ depends on Z^N. As an ultimate criterion of fit, Akaike (1981) suggested the use of the average information distance, or average entropy

$$E_{\hat{\theta}_N} I(\bar{f}_0(N, Z^N); \bar{f}(\hat{\theta}_N; N, Z^N)) \tag{7.92}$$

This is to be minimized with respect to both the model set and $\hat{\theta}_N$. As an unbiased estimate of the quantity (7.92), he suggested

$$\log \bar{f}(\hat{\theta}_N; N, Z^N) - \dim \theta \tag{7.93}$$

Calculations supporting this estimate will be given in Section 16.4.

The expression (7.93) used in (7.89) gives, with (7.72) and (7.73),

$$\hat{\theta}_{\text{AIC}}(Z^N) = \arg \min_{\theta} \left\{ \frac{1}{N} \sum_{t=1}^{N} \ell(\varepsilon(t, \theta), t, \theta) + \frac{\dim \theta}{N} \right\} \tag{7.94}$$

This is Akaike's information theoretic criterion (AIC). When applied to a given model structure, this estimate does not differ from the MLE in the same structure. The advantage with (7.94) is, however, that the minimization can be performed with respect to different model structures, thus allowing for a general identification theory. See Section 16.4 for a further discussion of this aspect.

An approach that is conceptually related to information measures is Rissanen's minimum description length (MDL) principle. This states that a model should be sought that allows the shortest possible code or description of the observed data. See Rissanen (1978, 1986). Within a given model structure, it gives estimates that coincide with the MLE. See also Section 16.4.

A Pragmatic Viewpoint

It is good and reassuring to know that general and sound basic principles, such as maximum likelihood, maximum entropy, and minimum information distance, lead to criteria of the kind (7.11). However, in the end we are faced with a sequence of figures that are to be compared with "guesses" produced by the model. It could then always be questioned whether a probabilistic framework and abstract principles are applicable, since we observe only a given sequence of data, and the framework relates to the thought experiment that the data collection can be repeated infinitely many times under "similar" conditions. It is thus an important feature that minimizing (7.11) makes sense, even without a probabilistic framework and without "alibis" provided by abstract principles.

7.5 CORRELATING PREDICTION ERRORS WITH PAST DATA

Ideally, the prediction error $\varepsilon(t,\theta)$ for a "good" model should be independent of past data Z^{t-1}. For one thing, this condition is inherent in a probabilistic model, such as (7.69). Another and more pragmatic way of seeing this condition is that if $\varepsilon(t,\theta)$ is correlated with Z^{t-1} then there was more information available in Z^{t-1} about $y(t)$ than picked up by $\hat{y}(t|\theta)$. The predictor is then not ideal. This leads to the characterization of a good model as one that produces prediction errors that are independent of past data.

A test if $\varepsilon(t,\theta)$ is independent of the whole (and increasing) data set Z^{t-1} would amount to testing whether all nonlinear transformations of $\varepsilon(t,\theta)$ are uncorrelated with all possible functions of Z^{t-1}. This is of course not feasible in practice.

Instead, we may select a certain finite-dimensional vector sequence $\{\zeta(t)\}$ derived from Z^{t-1} and demand a certain transformation of $\{\varepsilon(t,\theta)\}$ to be uncorrelated with this sequence. This would give

$$\frac{1}{N}\sum_{t=1}^{N}\zeta(t)\alpha(\varepsilon(t,\theta)) = 0 \tag{7.95}$$

and the θ-value that satisfies this equation would be the best estimate $\hat{\theta}_N$ based on the observed data. Here $\alpha(\varepsilon)$ is the chosen transformation of ε, and the typical choice would be $\alpha(\varepsilon) = \varepsilon$.

We may carry this idea into a somewhat higher degree of generality. In the first place, we could replace the prediction error with filtered versions as in (7.10).

Second, we obviously have considerable freedom in choosing the sequence $\zeta(t)$. It is quite possible that what appears to be the best choice of $\zeta(t)$ may depend on properties of the system. In such a case we would let $\zeta(t)$ depend on θ, and we have the following method:

Choose a linear filter $L(q)$ and let

$$\varepsilon_F(t,\theta) = L(q)\varepsilon(t,\theta) \tag{7.96a}$$

Choose a sequence of correlation vectors

$$\zeta(t,\theta) = \zeta(t,Z^{t-1},\theta) \tag{7.96b}$$

constructed from past data and, possibly, from θ. Choose a function $\alpha(\varepsilon)$. Then calculate

$$\hat{\theta}_N = \operatorname*{sol}_{\theta \in D_{\mathcal{M}}} [f_N(\theta,Z^N) = 0] \tag{7.96c}$$

$$f_N(\theta,Z^N) = \frac{1}{N}\sum_{t=1}^{N} \zeta(t,\theta)\alpha(\varepsilon_F(t,\theta)) \tag{7.96d}$$

Here we used the notation

$$\text{sol}[f(x) = 0] = \text{the solution(s) to the equation } f(x) = 0$$

Normally, the dimension of ζ would be chosen so that f_N is a d-dimensional vector (which means that ζ is $d \times p$ if the output is a p-vector). Then (7.96c) has as many equations as unknowns. In some cases it may be useful to consider an augmented correlation sequence ζ of higher dimension than d so that (7.96c) is an overdetermined set of equations, typically without any solution. Then the estimate is taken to be the value that minimizes some quadratic norm of f_N:

$$\hat{\theta}_N = \arg\min_{\theta \in D_{\mathcal{M}}} |f_N(\theta,Z^N)| \tag{7.97}$$

There are obviously formal links between these correlation approaches and the minimization approach of Section 7.2 (see, e.g., Problem 7D.6).

The procedure (7.96) is a conceptual method that takes different shapes, depending on which model structures it is applied to and on the particular choices of ζ. In the subsequent section we shall discuss the perhaps best known representatives of the family (7.96), the instrumental-variable methods. First, however, we shall discuss the pseudolinear regression models.

Pseudolinear Regressions

We found in Chapter 4 that a number of common prediction models could be written as

$$\hat{y}(t|\theta) = \varphi^T(t,\theta)\theta \tag{7.98}$$

[see (4.21) and (4.45)]. If the data vector $\varphi(t,\theta)$ does not depend on θ, this relationship would be a linear regression. From this the term pseudolinear regression for (7.98) is derived (Solo, 1978). For the model (7.98), the "pseudo-regression vector" $\varphi(t,\theta)$ contains relevant past data, partly reconstructed using the current model. It is thus reasonable to require from the model that the resulting prediction errors be uncorrelated with $\varphi(t,\theta)$. That is, we choose $\zeta(t,\theta) = \varphi(t,\theta)$ and $\alpha(\varepsilon) = \varepsilon$ in (7.96) and arrive at the estimate

$$\hat{\theta}_N^{\text{PLR}} = \text{sol}\left\{\frac{1}{N}\sum_{t=1}^{N}\varphi(t,\theta)[y(t) - \varphi^T(t,\theta)\theta] = 0\right\} \tag{7.99}$$

which we term the *PLR estimate*.

Models subject to (7.98) also lend themselves to a number of variants of (7.99), basically corresponding to replacing $\varphi(t,\theta)$ with vectors in which the "reconstructed" (θ-dependent) elements are determined in some other fashion. See Section 10.4.

7.6 INSTRUMENTAL-VARIABLE METHODS

Instrumental Variables

Consider again the linear regression model (7.31):

$$\hat{y}(t|\theta) = \varphi^T(t)\theta \tag{7.100}$$

Recall that this model contains several typical models of linear and nonlinear systems. The least-squares estimate of θ is given by (7.34) and can also be expressed as

$$\hat{\theta}_N^{\text{LS}} = \text{sol}\left\{\frac{1}{N}\sum_{t=1}^{N}\varphi(t)[y(t) - \varphi^T(t)\theta] = 0\right\} \tag{7.101}$$

An alternative interpretation of the LSE is consequently that it corresponds to (7.96) with $L(q) = 1$ and $\zeta(t,\theta) = \varphi(t)$.

Now suppose that the data actually can be described as in (7.37):

$$y(t) = \varphi^T(t)\theta_0 + v_0(t) \tag{7.102}$$

We then found in Section 7.3 that the LSE $\hat{\theta}_N$ will not tend to θ_0 in typical cases, the reason being correlation between $v_0(t)$ and $\varphi(t)$. Let us therefore try a general correlation vector $\zeta(t)$ in (7.101). Following general terminology in the system identification field, we call such an application of (7.96) to a linear regression an *instrumental-variable* method (IV). The elements of ζ are then called *instruments* or *instrumental variables*. This gives

$$\hat{\theta}_N^{\text{IV}} = \text{sol}\left\{\frac{1}{N}\sum_{t=1}^{N}\zeta(t)[y(t) - \varphi^T(t)\theta] = 0\right\} \tag{7.103}$$

or

$$\hat{\theta}_N^{IV} = \left[\frac{1}{N}\sum_{t=1}^{N}\zeta(t)\varphi^T(t)\right]^{-1}\frac{1}{N}\sum_{t=1}^{N}\zeta(t)y(t) \tag{7.104}$$

provided the indicated inverse exists. For $\hat{\theta}_N$ to tend to θ_0 for large N, we see from (7.103) that then $(1/N)\sum_{t=1}^{N}\zeta(t)v_0(t)$ should tend to zero. For the method (7.103) to be successfully applicable to the system (7.102), we would thus require the following properties of the instrumental variable $\zeta(t)$ (replacing sample means by expectation):

$$\bar{E}\zeta(t)\varphi^T(t) \text{ be nonsingular} \tag{7.105}$$

$$\bar{E}\zeta(t)v_0(t) = 0 \tag{7.106}$$

In words, we could say that the instruments must be correlated with the regression variables but uncorrelated with the noise. Let us now discuss possible choices of instruments that could be subject to (7.105) and (7.106).

Choices of Instruments

Suppose that (7.100) is an ARX model

$$y(t) + a_1 y(t-1) + \cdots + a_{n_a} y(t-n_a) = b_1 u(t-1) + \cdots + b_{n_b} u(t-n_b) + v(t) \tag{7.107}$$

Suppose also that the true description (7.102) corresponds to (7.107) with the coefficients indexed by "zero." A natural idea is to generate the instruments similarly to (7.107) so as to secure (7.105), but at the same time not let them be influenced by $\{v_0(t)\}$. This leads to

$$\zeta(t) = K(q)[-x(t-1) -x(t-2)\ldots -x(t-n_a)\; u(t-1)\ldots u(t-n_b)]^T \tag{7.108}$$

where K is a linear filter and $x(t)$ is generated from the input through a linear system

$$N(q)x(t) = M(q)u(t) \tag{7.109}$$

Here

$$N(q) = 1 + n_1 q^{-1} + \cdots + n_{n_n} q^{-n_n}$$

$$M(q) = m_0 + m_1 q^{-1} + \cdots + m_{n_m} q^{-n_m} \tag{7.110}$$

Most instruments used in practice are generated in this way. Obviously, $\zeta(t)$ is obtained from past inputs by linear filtering and can be written, conceptually, as

$$\zeta(t) = \zeta(t, u^{t-1}) \tag{7.111}$$

If the input is generated in *open loop* so that it does not depend on the noise $v_0(t)$ in the system, then clearly (7.106) holds. Since both the φ-vector and the ζ-vector are generated from the same input sequence (φ contains in addition effects from v_0), it

might be expected that (7.105) should hold "in general." We shall return to this question in Section 8.6.

A simple and appealing choice of instruments is to first apply the LS method to (7.107) and then use the LS-estimated model for N and M in (7.109). The instruments are then chosen as in (7.108) with $K(q) = 1$. Systems operating in closed loop and systems without inputs call for other ideas. See Problem 7G.3 for some suggestions.

As outlined in Problem 7D.5, the use of the instrumental vector (7.108) to (7.110) is equivalent to the vector

$$\zeta^*(t) = \frac{K(q)}{N(q)}[u(t-1)\,u(t-2)\ldots u(t-n_a-n_b)]^T \tag{7.112}$$

The IV estimate $\hat{\theta}_N^{IV}$ in (7.104) is thus the same for ζ^* as for ζ in (7.108) and does not, for example, depend on the filter M in (7.110).

Model-dependent Instruments (*)

The quality of the estimate $\hat{\theta}_N^{IV}$ will depend on the choice of $\zeta(t)$. In Section 9.5 we shall derive general expressions for the asymptotic covariance of $\hat{\theta}_N^{IV}$ and examine them further in Section 15.3. It then turns out that it may be desirable to choose the filter in (7.109) equal to those of the true system: $N(q) = A_0(q)$; $M(q) = B_0(q)$. These are clearly not known, but we may let the instruments depend on the parameters in the obvious way:

$$\zeta(t,\theta) = K(q)[-x(t-1,\theta)\ldots -x(t-n_a,\theta)\,u(t-1)\ldots u(t-n_b)]^T$$
$$A(q)x(t,\theta) = B(q)u(t) \tag{7.113}$$

In general, we could write the generation of $\zeta(t,\theta)$:

$$\zeta(t,\theta) = K_u(q,\theta)u(t) \tag{7.114}$$

where $K_u(q,\theta)$ is a d-dimensional column vector of linear filters.

Including a prefilter (7.96a) and a "shaping" function $\alpha(\cdot)$ for the prediction errors, the IV method could be summarized as follows:

$$\varepsilon_F(t,\theta) = L(q)[y(t) - \varphi^T(t)\theta] \tag{7.115a}$$

$$\hat{\theta}_N^{IV} = \operatorname*{sol}_{\theta \in D_M}[f_N(\theta, Z^N) = 0] \tag{7.115b}$$

where

$$f_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N}\zeta(t,\theta)\alpha(\varepsilon_F(t,\theta)) \tag{7.115c}$$

$$\zeta(t,\theta) = K_u(q,\theta)u(t) \tag{7.115d}$$

Extended IV Methods (*)

So far in this section the dimension of ζ has been equal to dim θ. We may also work with augmented instrumental variable vectors with dimension dim $\zeta > d$. The resulting method, corresponding to (7.96) and (7.97), will be called an *extended IV method* and takes the form

$$\hat{\theta}_N^{\text{EIV}} = \arg\min_{\theta} \left| \frac{1}{N} \sum_{t=1}^{N} \zeta(t,\theta)\alpha(\varepsilon_F(t,\theta)) \right|_Q^2 \tag{7.116}$$

The subscript Q denotes Q-norm:

$$|x|_Q^2 = x^T Q x \tag{7.117}$$

In case ζ does not depend on θ and $\alpha(\varepsilon) = \varepsilon$, (7.116) can be solved explicitly. See Problem 7D.7.

Frequency-domain Interpretation (*)

Quite analogously to (7.20) to (7.25) in the prediction error case, the criterion (7.115c) can be expressed in the frequency domain using Parseval's relationship. We then assume that $\alpha(\varepsilon) = \varepsilon$ and that a linear generation of the instruments as in (7.114) is used. This gives

$$f_N(\theta, Z^N) \approx \frac{1}{2\pi} \int_{-\pi}^{\pi} [\hat{G}_N(e^{i\omega}) - G(e^{i\omega},\theta)] |U_N(\omega)|^2 \cdot A(e^{i\omega},\theta) \cdot L(e^{i\omega}) \cdot K_u(e^{-i\omega},\theta) d\omega \tag{7.118}$$

Here $A(q,\theta)$ is the A-polynomial that corresponds to θ in the model (7.107).

Multivariable Case (*)

Suppose now that the output is p-dimensional. Then the instrument $\zeta(t)$ is a $d \times p$ matrix. A linear generation of $\zeta(t,\theta)$ could still be written as (7.114), with the interpretation that the ith column of $\zeta(t,\theta)$ is given by

$$\zeta^{(i)}(t,\theta) = K_u^{(i)}(q,\theta)u(t) \tag{7.119}$$

where $K_u^{(i)}(q,\theta)$ is a $d \times m$ matrix filter. [$K_u(q,\theta)$ in (7.114) is thus a tensor, a "three-index entity"]. With $\alpha(\varepsilon)$ being a function from $\mathbf{R}^p$ to $\mathbf{R}^p$ and $L(q)$ a $p \times p$ matrix filter, the IV method is still given by (7.115).

7.7 SUMMARY

There are several ways to fit models in a given set to observed data. In this chapter we have pointed out two general procedures. Both deal with the sequence of prediction errors $\{\varepsilon(t,\theta)\}$ computed from the respective models using the observed data, and both could be said to aim at making this sequence "small."

The *prediction-error identification approach* (PEM) was defined by (7.10) to (7.12):

$$\hat{\theta}_N = \arg\min_{\theta \in D_{\mathcal{M}}} V_N(\theta, Z^N)$$

$$V_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} \ell(\varepsilon(t,\theta),\theta,t) \tag{7.120}$$

It contains well-known procedures, such as the least-squares (LS) method and the maximum-likelihood (ML) method and is at the same time closely related to Bayesian maximum a posteriori (MAP) estimation and Akaike's information criterion (AIC).

The *correlation approach* was defined by (7.96):

$$\varepsilon_F(t,\theta) = L(q)\varepsilon(t,\theta)$$

$$\hat{\theta}_N = \operatorname*{sol}_{\theta \in D_{\mathcal{M}}} [f_N(\theta, Z^N) = 0]$$

$$f_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} \zeta(t,\theta)\alpha(\varepsilon_F(t,\theta)) \tag{7.121}$$

It contains the instrumental-variable (IV) technique, as well as several methods for rational transfer function models.

System identification has often been described as an area crowded with seemingly unrelated ad hoc methods and tricks. The list of names of available and suggested methods is no doubt a very long one. It is our purpose, however, with this chapter, as well as Chapters 8 to 11, to point out that the number of underlying basic ideas is really quite small, and that it indeed is quite possible to orient oneself in the area of system identification with these basic ideas as a starting point.

7.8 BIBLIOGRAPHY

The parameter estimation methods described in this chapter all go back to basic methods in statistics. For general texts, we refer to Cramér (1946), Rao (1973), and Lindgren (1976).

Section 7.2: The term "prediction-error methods" was perhaps first coined in Ljung (1974), but it had long been realized that the common methods of system identification had aimed at making the prediction error small. From an operational point of view, the criterion (7.120) can be viewed as a nonlinear regression method. See, for example, Jennrich (1969) and Hannan (1971b). Various norms have been discussed (see Section 15.2). The ℓ_∞-norm (related to "unknown-but-bounded" disturbances) is discussed in Milanese and Tempo (1985) (cf. Problem 7G.7) and Fogel and Huang (1982). See also Schweppe (1973). The frequency-domain expressions for the prediction-error criteria go back to Whittle (1951), who dealt with the

input-free case. Among many references, we may mention Hannan (1970), Chapter VI, for a detailed study (still with no input). Related formulas are also given by Solo (1978), Ljung and Glover (1981), and Wahlberg and Ljung (1986).

Section 7.3: The statistical roots of the least-squares method are examined in Appendix II. The application to times series has its origin in the work of Yule (1927) and Walker (1931), with a first asymptotic analysis by Mann and Wald (1943). The application to dynamic systems with an input was made independently by several authors, with an early comprehensive description and analysis by Åström (1968), to some extent reprinted in Åström and Eykhoff (1971). A good account of different variants of the LS method is given in Hsia (1977).

Section 7.4: Whittle pioneered maximum likelihood methods for AR and ARMA models using frequency domain formulations, see, e.g., Whittle (1951). The principle of maximum likelihood was then applied to dynamical systems by Åström and Bohlin (1965) (ARMAX model structures) and Box and Jenkins (1970) [model structure (4.31)]. Since then a long list of articles has dealt with this approach. Åström (1980) may be singled out for a survey.

Frequency-domain variants or approximations of the likelihood function have been extensively used by Whittle (1951), Hannan (1970), and others. The Bayesian MAP approach is comprehensively treated in Peterka (1981a, b). The calculations leading to (7.84) were first given by Eaton (1967) and Akaike (1973). Entropy and information theoretic criteria have been discussed extensively by Akaike and Rissanen. We may single out Akaike (1974a, 1981) and Rissanen (1985, 1986) as recommended reading. A general reference on entropy and statistics is Kullback (1959). The use of cross-entropy for estimation is discussed in Shore and Johnson (1980) and is extensively treated in Musicus (1982).

Sections 7.5–7.6: The way to describe the "correlation approach" as presented here is novel, although the different methods are well known. The IV method was introduced into statistics and econometrics by Reiersøl (1941) and has been applied to many parameter estimation problems in statistics (see, e.g., Kendall and Stuart, 1961). Applications to dynamic systems in the control field have been pioneered by Wong and Polak (1967), Young (1965), and Mayne (1967). For applications to ARMA models see Stoica, Friedlander and Söderström (1986). A historical background is given by Young (1976). For recent comprehensive treatments, see Söderström and Stoica (1983) and Young (1984).

7.9 PROBLEMS

7G.1. *Input error and output error methods:* Consider a model structure

$$y(t) = G(q, \theta)u(t)$$

without a specified noise model. In the survey of Åström and Eykhoff (1971) identification methods that minimize "the output error"

$$\hat{\theta}_N = \arg\min \sum_{t=1}^{N} [y(t) - G(q,\theta)u(t)]^2$$

and the "input error"

$$\hat{\theta}_N = \arg\min \sum_{t=1}^{N} [u(t) - G^{-1}(q,\theta)y(t)]^2$$

are listed. Show that these methods are prediction error methods corresponding to particular choices of noise models $H(q,\theta)$.

7G.2. *Spectral analysis as a prediction error method:* Consider the model structure

$$G(e^{i\omega},\theta) = \sum_{k=1}^{n} (g_k^R + ig_k^I) W_\gamma(\omega - \omega_k)$$

$$\theta = [g_1^R\ g_1^I, \ldots, g_n^R\ g_n^I]^T$$

and let $H(e^{i\omega},\eta)$ be an arbitrary noise model parametrization. Let $\hat{\theta}_N$ be the prediction-error estimate obtained by minimization of (7.23) and (7.25):

$$\hat{\theta}_N = \arg\min_{\theta,\eta} \int_{-\pi}^{\pi} \frac{|\hat{\hat{G}}_N(e^{i\omega}) - G(e^{i\omega},\theta)|^2 |U_N(\omega)|^2}{|H(e^{i\omega},\eta)|^2} d\omega$$

(a) Consider the special case $H(e^{i\omega},\eta) \equiv 1$ and

$$W_\gamma(\omega) = \begin{cases} 1, & |\omega| \le \dfrac{\pi}{2n} \\ 0, & |\omega| > \dfrac{\pi}{2n} \end{cases}$$

$$\omega_k = \frac{(k-1)\pi}{n}$$

Show that $\hat{G}(e^{i\omega_k}, \hat{\theta}_N)$ is then given by (6.46).

(b) Assume, in the general case, that

$$H(e^{i\omega},\eta) \cdot W_\gamma(\omega - \omega_k) \approx H(e^{i\omega_k},\eta) \cdot W_\gamma(\omega - \omega_k)$$

$$G(e^{i\omega},\theta) \cdot W_\gamma(\omega - \omega_k) \approx G(e^{i\omega_k},\theta) \cdot W_\gamma(\omega - \omega_k)$$

Show that (6.46) then holds approximately.

7G.3. *Instruments for closed-loop systems:* Consider a system

$$A_0(q)y(t) = B_0(q)u(t) + v_0(t)$$

under the output feedback

$$u(t) = F_1(q)r(t) - F_2(q)y(t)$$

(a) Let $x(t)$ and $\zeta(t)$ be given by

$$N(q)x(t) = M(q)r(t)$$

$$\zeta(t) = K(q)[-x(t-1)\ldots -x(t-n_a)\ r(t-1)\ldots r(t-n_b)]^T$$

Show that (7.106) holds for these instruments, and verify that (7.105) holds for a simple first-order special case.

(b) Suppose that $v_0(t)$ is known to be a MA process of order s. Introduce the instruments

$$\zeta(t) = [-y(t-1-s)\ldots -y(t-n_a-s)\, u(t-1-s)\ldots u(t-n_b-s)]^T$$

Show the same results as under part (a). See also Söderström, Stoica and Trulsson (1987).

7G.4. Suppose $Y_N = [y(1),\ldots,y(N)]^T$ is a Gaussian N-dimensional random vector with zero mean and covariance matrix $R_N(\theta)$. Let

$$R_N(\theta) = L_N(\theta)\Lambda_N(\theta)L_N^T(\theta)$$

where $L_N(\theta)$ is lower triangular with 1's along the diagonal and $\Lambda_N(\theta)$ a diagonal matrix with $\lambda_\theta(t)$ as the t, t element. Let

$$E_N(\theta) = L_N^{-1}(\theta)Y_N$$

$$E_N(\theta) = [\varepsilon(1,\theta),\ldots,\varepsilon(N,\theta)]^T$$

Show that, if θ is a parameter to be estimated, then the negative log likelihood function when Y_N is observed is

$$\frac{N}{2}\log 2\pi + \frac{1}{2}\log\det R_N(\theta) + \frac{1}{2}Y_N^T R_N^{-1}(\theta)Y_N$$

Show also that this can be rewritten as

$$\frac{N}{2}\log 2\pi + \frac{1}{2}\sum_{t=1}^{N}\log\lambda_\theta(t) + \frac{1}{2}\sum_{t=1}^{N}\frac{\varepsilon^2(t,\theta)}{\lambda_\theta(t)}$$

where $\varepsilon(t,\theta)$ are independent, normal random variables with variances $\lambda_\theta(t)$. How does this relate to our calculations (7.69) to (7.75)?

7G.5. Let the two random vectors X and Y be jointly Gaussian with

$$EX = m_X; \qquad EY = m_Y$$

$$E(X-m_X)(X-m_X)^T = P_X \qquad E(Y-m_Y)(Y-m_Y)^T = P_Y$$

$$E(X-m_X)(Y-m_Y)^T = P_{XY}$$

Show that the conditional distribution of X given Y is

$$(X|Y) \in N(m_X + P_{XY}P_Y^{-1}(Y-m_Y), P_X - P_{XY}P_Y^{-1}P_{XY}^T)$$

7G.6. Consider the model structure

$$\begin{aligned} X &= F(\theta)W \\ Y &= H(\theta)X + E \end{aligned} \tag{7.122}$$

where W and E are two independent, Gaussian random vectors with zero mean values and unit covariance matrices. Note that state-space models like (4.81), without input, can be written in this form by forming $X^T = [x^T(1)\ x^T(2)\ldots x^T(N)]$ and $Y^T = [y(1)\,y(2)\ldots y(N)]$. Let

$$R(\theta) = I + H(\theta)F(\theta)F^T(\theta)H^T(\theta)$$

Show the following:

(a) The negative log likelihood function for θ, (ignoring θ-independent terms) when Y is observed is

$$V(\theta) = -\log p(Y|\theta) = \tfrac{1}{2} Y^T R^{-1}(\theta) Y + \tfrac{1}{2} \log \det R(\theta)$$

Let

$$\hat{\theta}_{\mathrm{ML}} = \arg \min_{\theta} V(\theta)$$

(cf. Problem 7G.4).

(b) Let the conditional expectation of X, given Y and θ be $\hat{X}^s(\theta)$. Show that

$$E(X|Y, \theta) = \hat{X}^s(\theta) = [F(\theta)F^T(\theta)H^T(\theta)]R^{-1}(\theta)Y \tag{7.123}$$

and that

$$-\log p(X|\theta, Y) = \tfrac{1}{2}(X - \hat{X}^s(\theta))^T S^{-1}(\theta)(X - \hat{X}^s(\theta)) + \tfrac{1}{2} \log \det S(\theta)$$

$$S(\theta) = F(\theta)F^T(\theta) - F(\theta)F^T(\theta)H^T(\theta)R^{-1}(\theta)H(\theta)F(\theta)F^T(\theta) \tag{7.124}$$

(cf. Problem 7G.5) [$\hat{X}^s(\theta)$ gives the *smoothed* state estimate for the underlying state space model, see Anderson and Moore (1979)].

(c) Assume that the prior distribution of θ is flat. ($p(\theta) \approx$ independent of θ). Then show that the joint MAP estimate (7.65) of θ and X given Y,

$$(\hat{\theta}^s_{\mathrm{MAP}}, \hat{X}^s_{\mathrm{MAP}}) = \arg \max_{\theta, X} p(\theta, X|Y)$$

is given by

$$\arg \min_{X, \theta} [-\log p(Y, X|\theta)]$$

where

$$-\log p(Y, X|\theta) = \tfrac{1}{2}|Y - H(\theta)X|^2 + \tfrac{1}{2}|F^{-1}(\theta)X|^2 + \log \det F(\theta) \tag{7.125}$$

(d) Show that the value of X that minimizes (7.125) for fixed Y and θ is $\hat{X}^s(\theta)$, defined by (7.123). Hence

$$\hat{\theta}^s_{\mathrm{MAP}} = \arg \min_{\theta} \{\tfrac{1}{2}|Y - H(\theta)\hat{X}^s(\theta)|^2 + \tfrac{1}{2}|F^{-1}(\theta)\hat{X}^s(\theta)|^2 + \log \det F(\theta)\}$$

$$\hat{X}^s_{\mathrm{MAP}} = \hat{X}^s(\hat{\theta}^s_{\mathrm{MAP}})$$

(e) Establish that

$$-\log p(Y|\theta) = -\log p(Y, X|\theta) + \log p(X|\theta, Y) \tag{7.126}$$

(f) Establish that

$$-\log p(Y|\theta) = [\tfrac{1}{2}|Y - H(\theta)\hat{X}^s(\theta)|^2 + \tfrac{1}{2}|F^{-1}(\theta)\hat{X}^s(\theta)|^2 + \tfrac{1}{2} \log \det R(\theta)] \tag{7.127}$$

[Hint: Use the matrix identity (cf. (7.124))

$$S(\theta) = [F^{-T}(\theta)F^{-1}(\theta) + H^T(\theta)H(\theta)]^{-1}$$

and the determinant identity

$$\det(I_r + AB) = \det(I_s + BA)$$

for A and B $r \times s$ and $s \times r$ matrices and I_r the $r \times r$ identity matrix.]

(g) Conclude that $\hat{\theta}_{ML} \neq \hat{\theta}^{s}_{MAP}$ in general.

Remark: The problem illustrates the relationships among various expressions for the likelihood function, the smoothing problem, and MAP-estimates. "Log likelihood functions" of the kind (7.125) have been discussed, e.g., in Sage and Melsa (1971) and Schweppe (1973), Section 14.3.2.

7G.7. Consider the linear regression structure

$$y(t) = \varphi^T(t)\theta + v(t)$$

Based on the theory of optimal algorithms for operator approximation, (Traub and Wozniakowski, 1980), Milanese and Tempo (1985) and Milanese et al (1986) have suggested the following estimate:

For given δ, y_N and $\{\varphi(t)\}_1^N$, define the set

$$A_\delta = \{\theta \,|\, |y(t) - \varphi^T(t)\theta| < \delta \qquad \text{all } t = 1, \ldots, N\}$$

Assuming A_δ to be bounded and non-empty define its "center" $\theta_c(A_\delta)$ as follows: The ith component is

$$[\theta_c(A_\delta)]^{(i)} = \tfrac{1}{2}[\sup_{\theta \in A_\delta} \theta^{(i)} + \inf_{\theta \in A_\delta} \theta^{(i)}]$$

(superscript (i) denoting i : th component). The estimate $\hat{\theta}^{\delta}_N$ is then taken as $\theta_c(A_\delta)$.

(a) Suppose that $\dim \theta = 1$. Prove that $\hat{\theta}^{\delta}_N$ is independent of δ, as long as A_δ is nonempty and bounded.

(b) When $\dim \theta > 1$, $\hat{\theta}^{\delta}_N$ may in general depend on δ. Suppose that as δ decreases to a value δ^*, A_δ reduces to a singleton

$$A_{\delta^*} = \{\theta^\dagger\}$$

Then clearly $\hat{\theta}^{\delta^*}_N = \theta^\dagger$. Show that

$$\hat{\theta}^{\delta^*}_N = \arg \min_\theta \max_t |y(t) - \varphi^T(t)\theta|$$

This "optimal estimate" thus corresponds to the prediction error estimate (7.12) with the ℓ_∞-norm

$$\ell_\infty(\varepsilon^N(\cdot\,, \theta)) = \max_t |\varepsilon(t, \theta)|$$

This in turn can be seen as the limit as $p \to \infty$ of the criterion functions

$$\ell(\varepsilon) = |\varepsilon|^p$$

in (7.11).

7E.1. *Estimating the AR Part of an ARMA model:* Consider the ARMA model

$$A(q)y(t) = C(q)e(t)$$

with orders n_a and n_c, respectively. A method to estimate the AR part has been given as follows. Let

$$\hat{R}^N_y(\tau) = \frac{1}{N} \sum_{t=\tau}^{N} y(t)y(t - \tau)$$

Then solve for $\hat{a}_i^N$ from

$$\hat{R}_y^N(\tau) + a_1 \hat{R}_y^N(\tau - 1) + \cdots + a_{n_a} \hat{R}_y^N(\tau - n_a) = 0$$

$$\tau = n_c + 1, n_c + 2, \ldots, n_c + n_a$$

Show that this (essentially) is an application of the IV method using specific instruments. Which ones? (See Cadzow, 1980, and Stoica, Söderström, and Friedlander, 1985.)

7E.2. *Sinusoids in noise:* Consider a sinusoid measured in white Gaussian noise:

$$y(t) = \alpha e^{i\omega t} + e(t)$$

For simplicity we use complex algebra. The constant α is thus complex-valued. The amplitude, phase and frequency are unknown: $\theta = (\alpha, \omega)$. The predictor thus is

$$\hat{y}(t|\theta) = \alpha e^{i\omega t}$$

If $e(t)$ has variance 1 (real and imaginary parts independent), the likelihood function gives the prediction-error criterion:

$$V_N(\theta, Z^N) = \frac{1}{2} \sum_{t=1}^{N} |y(t) - \hat{y}(t|\theta)|^2$$

Show that the MLE

$$\hat{\theta}_N = \begin{bmatrix} \hat{\alpha}_N \\ \hat{\omega}_N \end{bmatrix} = \arg\min_{\theta} V_N(\theta, Z^N)$$

obeys

$$\hat{\omega}_N = \arg\max_{\omega} |Y_N(\omega)|^2$$

where $Y_N(\omega)$ is the Fourier transform (2.37) of $y(t)$.

7E.3. *Error-in-variables models:* Econometric models often include disturbances both on inputs and outputs (compare our comment in Section 2.1 on Figure 2.2). Consider the model in Figure 7.1. The true inputs and outputs are thus s and x, while we measure u and y. In a first-order case, we have

$$x(t) + ax(t-1) = bs(t-1)$$

$$y(t) = x(t) + e(t)$$

$$u(t) = s(t) + w(t)$$

Suppose that w and e are independent white noises with unknown variances. Discuss how a, b, and these variances can be estimated using measurements of y and u.

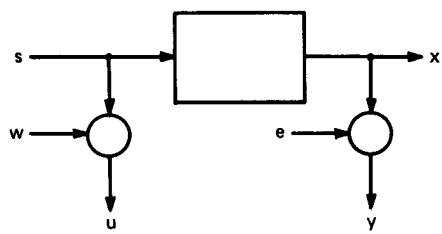

Figure 7.1 An error-in-variables model.

[*Remark:* With the assumption that the color of the noises are known, the problem is relatively simple. Without this assumption ("prejudice"), the problem is not yet solved. See Kalman (1983), B.D.O. Anderson (1985), and Söderström (1981)].

7E.4. Consider a probabilistic model, implicitly given in the state-space form

$$x(t+1) = ax(t) + w(t)$$

$$y(t) = x(t) + e(t) \tag{7.122a}$$

where $\{w(t)\}$ and $\{e(t)\}$ are assumed to be independent, white Gaussian noises, with variances

$$Ew^2(t) = r_1$$

$$Ee^2(t) = 1 \quad \text{(assumed known)} \tag{7.122b}$$

Let the parameter vector be

$$\theta = \begin{bmatrix} a \\ r_1 \end{bmatrix} \tag{7.122c}$$

Assume initial conditions for $x(0)$ (mean and variance) such that the prediction $\hat{y}(t|\theta)$ becomes a stationary process for each θ (i.e., so that the steady-state Kalman filter can be used). Determine the log-likelihood function for this problem. Compare with the log-likelihood function for a directly parametrized innovations representation model (4.88).

7E.5. Consider the nonlinear model structure of Problem 5E.1. Discuss how the LS, ML, IV, and PLR methods can be applied to this structure. (Reference: Fnaiech and Ljung, 1986).

7E.6. Consider the model structure

$$y(t) = \varphi^T(t)\theta + v(t)$$

where the regression vector $\varphi(t)$ can only be measured with noise:

$$\eta(t) = \varphi(t) + w(t)$$

The noises $\{w(t)\}$ and $\{v(t)\}$ may be nonwhite and mutually correlated. Suppose a vector $\zeta(t)$ is known that is uncorrelated with $\{v(t)\}$ and $\{w(t)\}$ but correlated with $\varphi(t)$. Suggest how to estimate θ from $y(t)$, $\eta(t)$, and $\zeta(t), t = 1, \ldots, N$.

7E.7. Suppose in (7.74) and (7.75) that λ does not depend on θ. Determine $\hat{\lambda}_N$.

7E.8. Consider the model structure

$$\hat{y}(t|\theta) = -ay(t-1) + bu(t-1)$$

and assume that the true system is given by

$$y(t) - 0.9y(t-1) = u(t-1) + e_0(t)$$

where $\{e_0(t)\}$ is white noise of unit variance. Determine the Cramér–Rao bound for the estimation of a and b. How does it depend on the properties of u?

7T.1. Suppose that a true description of a certain system is given by

$$y(t) + a_1^0 y(t-1) + \cdots + a_{n_a}^0 y(t-n_a) = b_1^0 u(t-1) + \cdots + b_{n_b}^0 u(t-n_b) + v_0(t)$$

for a stationary process $\{v_0(t)\}$ independent of the input. Let $\varphi(t)$ be defined, as usual, by (7.32), and let $\tilde{\varphi}(t)$ be given by

$$\tilde{\varphi}(t) = [-y_0(t-1) \dots -y_0(t-n_a)\, u(t-1) \dots u(t-n_b)]^T$$

where

$$y_0(t) + a_1^0 y_0(t-1) + \cdots + a_{n_a}^0 y_0(t-n_a) = b_1^0 u(t-1) + \cdots + b_{n_b}^0 u(t-n_b)$$

Prove that for any vector of instrumental variables of the general kind (7.108) we have

$$E\zeta(t)\varphi^T(t) = E\zeta(t)\tilde{\varphi}^T(t)$$

7D.1. Consider the ARX structure (4.7) where one parameter, say b_1, is known to have a certain value b_1^*. Show that the associated predictor can be written as

$$\hat{y}(t|\theta) = \theta^T \varphi(t) + \mu(t)$$

with proper definitions of θ, φ, and μ (φ and μ to be known variables at time t). Derive the LS estimate and the IV estimate for this model.

7D.2. Let A be a given, positive symmetric definite matrix and let B and C be given matrices. Establish that

$$\theta^T A\theta - \theta^T B - B^T\theta + C = [\theta - A^{-1}B]^T A[\theta - A^{-1}B] + C - B^T A^{-1} B$$
$$\geq C - B^T A^{-1} B$$

and use this result to prove all the expressions for the LSE in Section 7.3 [(7.34), (7.42), (7.44), and (7.47)]. The matrix inequality $D \geq B$ is to be interpreted as "$D - B$ is a positive semidefinite matrix."

Hint: For (7.47), rewrite (7.46) as

$$V_N(\theta, Z^N) = \operatorname{tr} \frac{1}{N} \sum_{t=1}^{N} [y(t) - \theta^T \varphi(t)][y(t) - \theta^T \varphi(t)]^T$$

7D.3. Let Σ be an invertible square $p \times p$ matrix with elements σ_{ij}. Prove the differentiation formula

$$\frac{\partial}{\partial \sigma_{ij}} \det \Sigma = \det[\Sigma] \cdot \mu_{ij}$$

where μ_{ij} is the i, j element of Σ^{-1}. [*Hint:* Use $\det(I + \varepsilon A) = 1 + \varepsilon \operatorname{tr} A +$ higher-order terms in ε]. Use the result to prove (7.82) and (7.83).

7D.4. Show that the two instrumental variable vectors, of dimension d, $\zeta_1(t)$ and $\zeta_2(t)$, where $\rho_1(t) = T\rho_2(t)$ with T invertible, give the same estimate $\hat{\theta}_N^{\text{IV}}$ in (7.104).

7D.5. Show that if two variables x and u are associated as in (7.109) and (7.110) then we can write

$$\begin{bmatrix} -x(t-1) \\ \vdots \\ -x(t-n_n) \\ u(t-1) \\ \vdots \\ u(t-n_m) \end{bmatrix} = S(-M, N) \cdot \frac{1}{N(q)} \begin{bmatrix} u(t-1) \\ u(t-2) \\ \vdots \\ u(t-n_n-n_m) \end{bmatrix}$$

for an $(n_n + n_m) \times (n_n + n_m)$ matrix

$$S(-M, N) = \begin{bmatrix} -m_0 & -m_1 & \dots & -m_{n_m} & 0 & \dots & 0 \\ 0 & -m_0 & \dots & -m_{n_m-1} & -m_{n_m} & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & -m_0 & -m_1 & \dots & -m_{n_m} \\ 1 & n_1 & \dots & n_{n_n} & 0 & \dots & 0 \\ 0 & 1 & \dots & n_{n_n-1} & n_{n_n} & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & n_1 & \dots & n_{n_n} \end{bmatrix}$$

Such a matrix is called a *Sylvester matrix* (see, e.g., Kailath, 1980), and it will be nonsingular if and only if the polynomials in (7.110) have no common factor. Use this result to prove that the instruments (7.112) give the same IV estimate as the instruments (7.108). Reference: Söderström and Stoica (1983).

7D.6. Show that the prediction-error estimate obtained from (7.11) and (7.12) can also be seen as a correlation estimate (7.96) for a particular choice of L, ζ, and α.

7D.7. Give an explicit expression for the estimate $\hat{\theta}_N^{\text{EIV}}$ in (7.116) in the case ζ does not depend on θ, and $\alpha(\varepsilon) = \varepsilon$.

7D.8. Consider the symmetric matrix

$$H = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$$

Show that if $H \geq 0$, then

$$A - BC^{-1}B^T \geq 0.$$

Hint: Consider xHx^T for

$$x = [x_1 \ -x_1 BC^{-1}]$$

with x_1 arbitrary.

7S.1. Write a MACRO

$$\text{TH} = \text{LS}(y, u, na, nb, nk)$$

that computes the LS estimate (7.34) for the ARX model (4.7):

$$y(t) + a_1 y(t-1) + \cdots + a_{na} y(t - na) = b_{nk} u(t - nk) + \cdots + b_{nk+nb-1} u(t - nk + 1 - nb) + e(t)$$

Let the result TH be consistent with the format of Problem 4S.1 and let λ in TH be $2 \cdot V_N(\hat{\theta}_N, Z^N)$.

7S.2. Write a MACRO

$$\text{TH} = \text{IV}(y, u, na, nb, nk, N, M)$$

that computes the IV estimate (7.104), (7.108), (7.109) $[K(q) = 1]$ for the same model as in Problem 7S.1.

7C.1. Simulate the system

$$y(t) = \frac{q^{-1} + 0.5q^{-2}}{1 - 1.5q^{-1} + 0.7q^{-2}} u(t) + e_0(t)$$

over N samples, with $\{u(t)\}$ being a random binary ± 1 signal and $\{e_0(t)\}$ white Gaussian noise with variance 1. Use the obtained data record $[Z^N = (y^N, u^N)]$ to identify the system in the following ways:

(a) Determine the spectral analysis estimate using various values of γ (Problem 6S.1).

(b) Determine the LS estimate for various model orders, and compare the corresponding Bode plots with the spectral analysis estimate (and the true Bode plot). Why does the fit become better as the model orders increase?

(c) Determine the IV estimates for instruments generated as in (7.108) to (7.110) with $n_a = n_b = 2$ and

$c1$: $K(q) = 1, \quad N(q) = 1, \quad M(q) = q^{-2}$

$c2$: $K(q) = 1, \quad N, M$ being the LS estimates of A and B

$c3$: $K(q) = \dfrac{1}{1 - 1.5q^{-1} + 0.7q^{-2}},$

$N(q) = 1 - 1.5q^{-1} + 0.7q^{-2}$

$M(q) = q^{-1} + 0.5q^{-2}$

Try $n_a = n_b = 3$. What happens?

Try $N = 100$ and $N = 400$ and some different realizations to get a feel for the variability of the results.

7C.2. Simulate the system

$$y(t) - 0.9y(t-1) = u(t-1) + 0.5u(t-2) + e_0(t)$$

with u and e_0 as in Problem 7C.1. Let $N = 100$. Consider the two model structures

$$\mathcal{M}_1: \quad y(t) + ay(t-1) = u(t-1) + bu(t-2) + e(t)$$

$$\mathcal{M}_2: \quad y(t) = \frac{q^{-1} + bq^{-2}}{1 + fq^{-1}}u(t) + e(t)$$

Apply the quadratic prediction-error approach to each of them and draw mesh plots of the corresponding criterion functions as functions of (a, b), and (f, b) respectively.

APPENDIX 7A: PROOF OF THE CRAMÉR–RAO INEQUALITY

The assumption $E\hat{\theta}(y^N) = \theta_0$ can be written

$$\theta_0 = \int_{\mathbf{R}^N} \hat{\theta}(x^N) f_y(\theta_0, x^N) dx^N \tag{7A.1}$$

By definition we also have

$$1 = \int_{\mathbf{R}^N} f_y(\theta_0, x^N) dx^N \tag{7A.2}$$

Differentiating these two expressions with respect to θ_0 gives

$$
\begin{aligned}
I &= \int_{\mathbf{R}^N} \hat{\theta}(x^N)\left[\frac{d}{d\theta_0} f_y(\theta_0, x^N)\right]^T dx^N \\
&= \int_{\mathbf{R}^N} \hat{\theta}(x^N)\left[\frac{d}{d\theta_0} \log f_y(\theta_0, x^N)\right]^T f_y(\theta_0, x^N) dx^N \\
&= E\hat{\theta}(y^N)\left[\frac{d}{d\theta_0} \log f_y(\theta_0, y^N)\right]^T
\end{aligned} \tag{7A.3}
$$

(I is the $d \times d$ unit matrix) and

$$
\begin{aligned}
0 &= \int_{\mathbf{R}^N} \left[\frac{d}{d\theta_0} f_y(\theta_0, x^N)\right]^T dx^N = \int_{\mathbf{R}^N} \left[\frac{d}{d\theta_0} \log f_y(\theta_0, x^N)\right]^T f_y(\theta_0, x^N) dx^N \\
&= E\left[\frac{d}{d\theta_0} \log f_y(\theta_0, y^N)\right]^T
\end{aligned} \tag{7A.4}
$$

Expectation in these two expressions are hence w.r.t. y^N.

Now multiply (7A.4) by θ_0 and subtract it from (7A.3). This gives

$$
E[\hat{\theta}(y^N) - \theta_0]\left[\frac{d}{d\theta_0} \log f_y(\theta_0, y^N)\right]^T = I \tag{7A.5}
$$

Now denote

$$
\alpha = \hat{\theta}(y^N) - \theta_0, \qquad \beta = \frac{d}{d\theta_0} \log f_y(\theta_0, y^N) \tag{7A.6}
$$

(both d-dimensional column vectors) so that

$$
E\alpha\beta^T = I \tag{7A.7}
$$

Hence

$$
E\begin{bmatrix}\alpha \\ \beta\end{bmatrix}\begin{bmatrix}\alpha \\ \beta\end{bmatrix}^T = \begin{bmatrix} E\,\alpha\alpha^T & I \\ I & E\,\beta\beta^T \end{bmatrix} \geq 0
$$

where the positive semidefiniteness follows by construction. Hence Problem 7D.8 proves that

$$
E\alpha\alpha^T \geq [E\beta\beta^T]^{-1}
$$

which is (7.67). It only remains to prove the equality in (7.68). Differentiating the transpose of (7A.4) gives

$$
\begin{aligned}
0 = &\int_{\mathbf{R}^N} \left[\frac{d^2}{d\theta_0^2} \log f_y(\theta_0, x^N)\right] f_y(\theta_0, x^N) dx^N \\
&+ \int_{\mathbf{R}^N} \left[\frac{d}{d\theta_0} \log f_y(\theta_0, x^N)\right]\left[\frac{d}{d\theta_0} \log f_y(\theta_0, x^N)\right]^T f_y(\theta_0, x^N) dx^N
\end{aligned}
$$

which gives (7.68)

8

CONVERGENCE AND CONSISTENCY

8.1 INTRODUCTION

In Chapter 7 we described a number of different methods to determine models from data. To use these methods in practice, we need insight into their properties: How well will the identified model describe the actual system? Are some identification methods better than others? How should the design variables associated with a certain method be chosen?

Such questions relate, from a formal point of view, to the mapping (7.7) from the data set Z^N to the parameter estimate $\hat{\theta}_N$:

$$Z^N \rightarrow \hat{\theta}_N \in D_{\mathcal{M}} \tag{8.1}$$

Questions about properties of this mapping can be answered basically in two ways:

1. Generate data Z^N with known characteristics. Apply the mapping (8.1) (corresponding to a particular identification method) and evaluate the properties of $\hat{\theta}_N$. This is known as *simulation* studies.
2. Assume certain properties of Z^N and try to calculate what the inherited properties of $\hat{\theta}_N$ are. This is known as *analysis*.

In this chapter we shall analyze the *convergence properties* of $\hat{\theta}_N$ as N tends to infinity. Since we will never encounter infinitely many data, such analysis has the character of a "thought experiment," and we must support it with some assump-

tions about a corresponding infinite data set Z^∞. There are some different possibilities for such assumptions (see Problem 8T.1). Here we shall adopt a stochastic framework for the observations, along the lines described in Chapter 2. We shall thus consider the data as realizations of a stochastic process with deterministic components. It might be worthwhile to contemplate what analysis under such assumptions actually amounts to. A probabilistic framework relates to the following questions: What would happen if I repeat the experiment? Should I then expect a very different result? Will the limit of $\hat{\theta}_N$ depend on the particular realization of the random variables? Even if the experiment is never repeated, it is clear that such questions are relevant for the confidence one should develop for the estimate, and this makes the analysis worthwhile. It is then another matter that the probabilistic framework that is set up to answer such questions may exist only in the mind of the analyzer and cannot be firmly tied to the real-world experiment.

It should also be remarked that a conventional stochastic description of disturbances is not without problems: For example, suppose we measure a distance with a crude measuring rod and describe the measurement error as a zero-mean random variable, which is independent of the error obtained when the experiment is repeated. This assumption implies, by the law of large numbers, that the distance can be determined with arbitrary accuracy, if only the measurements are repeated sufficiently many times. Clearly such a conclusion can be criticized from a practical point of view. Results from theoretical analysis must thus be interpreted with care when applied to a practical situation.

The question of how $\hat{\theta}_N$ behaves as N increases clearly relates to the question of how the corresponding criteria functions $V_N(\theta, Z^N)$ and $f_N(\theta, Z^N)$ behave. These are, with a stochastic framework, sums of random variables, and their convergence properties will be consequences of the law of large numbers. Our basic technical tool in this chapter will thus be Theorem 2B.1. In order not to conceal the basic ideas with too much technicalities, we shall only complete the proofs for linear, time-invariant models (such as those in Chapter 4) and quadratic criteria. The techniques and results, however, carry over also to more general cases.

The chapter is organized as follows. Assumptions about the infinite data set Z^∞ are given in Section 8.2. Convergence for prediction-error estimates is treated in Section 8.3. Consistency questions (i.e., whether the true system is retrieved in the limit) are discussed in Section 8.4. A frequency-domain characterization of the limit estimate is given in Section 8.5. In Section 8.6, the corresponding results are given for the correlation approach.

A Preview

In the chapter a general and natural result is derived: the estimate $\hat{\theta}_N$ obtained by the prediction-error method (7.120) will converge to the value that minimizes the average criterion $\bar{E}\ell(\varepsilon(t, \theta), \theta)$. Here $\bar{E}$ can heuristically be taken as averaging over time or ensembles (possible realizations) or both. The chapter deals both with the formal framework for establishing this "obvious" result and with characterizations

of the limit value of $\hat{\theta}_N$. The reluctant reader of theory should concentrate on understanding the main result, Eq. (8.29), and the frequency-domain characterization of the limit model in Section 8.5.

8.2 CONDITIONS ON THE DATA SET

The data set

$$Z^N = \{u(1), y(1), \ldots, u(N), y(N)\}$$

is the basic starting point. Analysis, we said, amounts to assuming certain properties about the data and computing the resulting properties of $\hat{\theta}_N$. Since the analysis of $\hat{\theta}_N$ will be carried out for $N \to \infty$, it is natural that the conditions on the data relate to the infinite set Z^∞. In this section we shall introduce such conditions, as well as some pertinent definitions.

A Technical Condition D1 (*)

We shall assume that the actual data are generated as depicted in Figure 8.1. The input u may be generated (partly) as output feedback or in open loop ($u = w$). The signal e_0 represents the disturbances that act on the process. [The subscript 0 distinguishes this "true" noise e_0 from the "dummy" noise e we have used in our model descriptions (7.3)]. The prime objective with condition D1 is to describe the closed-loop system in Figure 8.1 as a stable system so that the dependence between far apart data decays. The most restrictive condition is the assumed linearity (8.2). It can be traded for more general conditions, at the price of more complicated analysis. See Ljung (1978a), condition S3. For our analysis, we shall use the following technical assumptions:

D1: The data set Z^∞ is such that for some filters $\{d_i^{(i)}(k)\}$

$$\begin{aligned} y(t) &= \sum_{k=1}^{\infty} d_i^{(1)}(k) w(t-k) + \sum_{k=0}^{\infty} d_i^{(2)}(k) e_0(t-k) \\ u(t) &= \sum_{k=0}^{\infty} d_i^{(3)}(k) w(t-k) + \sum_{k=0}^{\infty} d_i^{(4)}(k) e_0(t-k) \end{aligned} \tag{8.2}$$

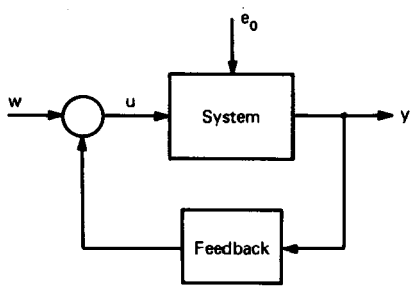

Figure 8.1 The data-generating configuration.

where

1. $\{w(t)\}$ is a bounded, deterministic, external input sequence. (8.3)
2. $\{e_0(t)\}$ is a sequence of independent random variables with zero mean values and bounded moments of order $4 + \delta$ for some $\delta > 0$. (8.4)

Moreover,

3. The family of filters $\{d_t^{(i)}(k)\}_{k=1}^{\infty}, i = 1\text{–}4; t = 1, 2, \ldots$ is uniformly stable. (8.5)
4. The signals $\{y(t)\}, \{u(t)\}$ are jointly quasi-stationary. (8.6)

Recall the definitions of stability (2.29) and quasi-stationarity (2.58) to (2.62). (Problem 2T.4 showed that uniform stability holds, even if the closed-loop system goes through "unstable transients.")

Remark. When we say that $\{w(t)\}$ is "deterministic," we simply mean that we regard it as a given sequence that (in contrast to e_0) can be reproduced if the experiment is repeated. The stochastic operators and qualifiers, such as E, w.p. 1, and AsN will thus average over the properties of $\{e_0(t)\}$ for the fixed sequence $\{w(t)\}$. Of course, this does not exclude that this particular sequence $\{w(t)\}$ actually is generated as a realization of a stochastic process, independent of the system disturbances. In that case it is sometimes convenient to let the expectation also average over the probabilistic properties of $\{w(t)\}$. We shall comment on how to do this below [Eq. (8.27)]. ■

A True System $\mathcal{S}$

We shall sometimes use a more specific assumption of a "true system":
S1: The data set Z^{∞} is generated according to

$$\mathcal{S}: \quad y(t) = G_0(q)u(t) + H_0(q)e_0(t) \tag{8.7}$$

where $\{e_0(t)\}$ is a sequence of independent random variables, with zero mean values, variances λ_0, and bounded moments of order $4 + \delta$, some $\delta > 0$, and $H_0(q)$ is an inversely stable, monic filter.

We thus denote the true system by $\mathcal{S}$. Given a model structure (4.4),

$$\mathcal{M}: \quad \{G(q, \theta), H(q, \theta) | \theta \in D_{\mathcal{M}}\} \tag{8.8}$$

it is natural to check whether the true system (8.7) belongs to the set defined by (8.8). We thus introduce

$$D_T(\mathcal{S}, \mathcal{M}) = \{\theta \in D_{\mathcal{M}} | G(e^{i\omega}, \theta) = G_0(e^{i\omega}); H(e^{i\omega}, \theta) = H_0(e^{i\omega}); -\pi \leq \omega \leq \pi\} \tag{8.9}$$

This set is nonempty precisely when the model structure admits an exact description of the true system. We write this also as

$$\mathcal{S} \in \mathcal{M} \tag{8.10}$$

Although such an assumption is not particularly realistic in practical applications, it yields a quite useful insight into properties of the estimated models.

When S1 holds, a more explicit version of conditions D1 can be given:

Lemma 8.1. Suppose that S1 holds, and the input is chosen as

$$u(t) = -F(q)y(t) + w(t)$$

such that

$$[1 + G_0(q)F(q)]^{-1}G_0(q), [1 + G_0(q)F(q)]^{-1}H_0(q), F(q)[1 + G_0(q)F(q)]^{-1}G_0(q)$$

and

$$F(q)[1 + G_0(q)F(q)]^{-1}H_0(q)$$

are stable filters and that $\{w(t)\}$ is quasi-stationary. Then condition D1 holds.

Proof. We have, for the closed-loop systems,

$$y(t) = [1 + G_0(q)F(q)]^{-1}G_0(q)w(t) + [1 + G_0(q)F(q)]^{-1}H_0(q)e_0(t) \tag{8.11}$$

The stability condition means that the filters in (8.11) are stable. Thus (8.6) follows from Theorem 2.2. Moreover, (8.2), and (8.5), are immediate from (8.11) and the stability assumption. ■

Information Content in the Data Set

The set Z^N is our source of information about the true system. This is to be fit to a model structure $\mathcal{M}$ of our choice. (The reader might at this point review Section 4.5, if necessary.) The structure $\mathcal{M}$ describes a set of models $\mathcal{M}^*$ within which the best one is sought for. Identifiability of model structures concerns the question whether different parameter vectors may describe the same model in the set $\mathcal{M}^*$. See Definitions 4.6 to 4.8. A related question is whether the data set Z^∞ allows us to distinguish between different models in the set. Recall that, according to Definition 4.1, a (linear time-invariant) model is given by a filter $W(q)$. We shall call a data set *informative* if it is capable of distinguishing between different models. We thus introduce the following concept:

Definition 8.1. A quasi-stationary data set Z^∞ is *informative enough with respect to the model set* $\mathcal{M}^*$ if, for any two models $W_1(q)$ and $W_2(q)$ in the set,

$$\bar{E}[(W_1(q) - W_2(q))z(t)]^2 = 0 \tag{8.12a}$$

implies that $W_1(e^{i\omega}) \equiv W_2(e^{i\omega})$ almost all ω. ■

We note that with

$$W_1(q) - W_2(q) = [\Delta W_u(q) \;\; \Delta W_y(q)]$$

(8.12a) can be written

$$\bar{E}[\Delta W_u(q)u(t) + \Delta W_y(q)y(t)]^2 = 0 \tag{8.12b}$$

Note that the limit in (8.12) exists in view of (8.6) and Theorem 2.2. Recall also (4.109) and the definition of equality of models, (4.113).

Definition 8.2. A quasi-stationary data set Z^∞ is *informative* if it is informative enough with respect to the model set $\mathcal{L}^*$, consisting of all linear, time-invariant models. ■

The concept of informative data sets is very closely related to concepts of "persistently exciting" inputs, "general enough" inputs, and so on. We shall discuss the concept in detail in Chapter 14 in connection with experiment design. Here we give an immediate consequence of Definition 8.2.

Theorem 8.1. A quasi-stationary data set Z^∞ is informative if the spectrum matrix for $z(t) = [u(t)\; y(t)]^T$ is strictly positive definite for all ω.

Proof. Consider (8.12) for arbitrary linear models W_1 and W_2. Denote $W_1(q) - W_2(q) = \tilde{W}(q)$. Then applying Theorem 2.2 to (8.12) gives

$$0 = \int_{-\pi}^{\pi} \tilde{W}(e^{i\omega})\Phi_z(\omega)\tilde{W}^T(e^{-i\omega})d\omega$$

where

$$\Phi_z(\omega) = \begin{bmatrix} \Phi_u(\omega) & \Phi_{uy}(\omega) \\ \Phi_{yu}(\omega) & \Phi_y(\omega) \end{bmatrix} \tag{8.13}$$

Since $\Phi_z(\omega)$ is positive definite, this implies that $\tilde{W}(e^{i\omega}) \equiv 0$ almost everywhere, which proves the theorem. ■

Some Additional Concepts and Notations (*)

In Definition 4.3 we defined a model structure as a differentiable mapping, such that the predictors and their gradients were stable for each $\theta \in D_{\mathcal{M}}$. To facilitate the analysis, we now strengthen this condition.

Definition 8.3. A model structure $\mathcal{M}$ is said to be *uniformly stable* if the family of filters $\{W(q,\theta),\ \Psi(q,\theta)$ and $(d/d\theta)\Psi(q,\theta);\ \theta \in D_{\mathcal{M}}\}$ is uniformly stable. ■

Analogous to (4.106), we shall, when S1 holds, define

$$\chi_0(t) = \begin{bmatrix} u(t) \\ e_0(t) \end{bmatrix} \tag{8.14a}$$

and

$$T_0(q) = [G_0(q)\ H_0(q)] \tag{8.14b}$$

The system (8.7) can thus be written

$$y(t) = T_0(q)\chi_0(t)$$

The difference will be denoted

$$\tilde{T}(q,\theta) = T_0(q) - T(q,\theta) = [\tilde{G}(q,\theta)\ \tilde{H}(q,\theta)] \tag{8.15}$$

8.3 PREDICTION-ERROR APPROACH

Basic Result

The prediction-error estimate is defined by (7.12)

$$\hat{\theta}_N = \arg\min_{\theta \in D_{\mathcal{M}}} V_N(\theta, Z^N) \tag{8.16}$$

To determine the limit to which $\hat{\theta}_N$ converges as N tends to infinity is obviously related to the limit properties of the function $V_N(\theta, Z^N)$. For a quadratic criterion and a linear, uniformly stable model structure $\mathcal{M}$, we have

$$V_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} \frac{1}{2}\varepsilon^2(t,\theta) \tag{8.17}$$

and, using (7.2),

$$\varepsilon(t,\theta) = [1 - W_y(q,\theta)]y(t) - W_u(q,\theta)u(t) \tag{8.18}$$

Under assumption D1 we can replace $y(t)$ and $u(t)$ in the preceding expression by (8.2), which gives

$$\varepsilon(t,\theta) = \sum_{k=1}^{\infty} d_i^{(5)}(k;\theta)w(t-k) + \sum_{k=0}^{\infty} d_i^{(6)}(k;\theta)\, e_0(t-k) \tag{8.19}$$

Now the filters in (8.18) are uniformly (in θ) stable since $\mathcal{M}$ is uniformly stable. Under assumption (8.5) the filters in (8.2) are uniformly (in t) stable. Hence the cascaded filters $\{d_i^{(i)}(k;\theta)\}$, $i = 5, 6$, in (8.19) are also uniformly (in both θ and t) stable (see Problem 8D.2). That is,

$$|d_i^{(i)}(k;\theta)| \le \beta_k, \ \forall t, \forall\theta \in D_{\mathcal{M}}, i = 5, 6 \qquad \sum_{1}^{\infty} \beta_k < \infty \tag{8.20}$$

Finally, under assumption (8.6), Theorem 2.2 implies that $\{\varepsilon(t,\theta)\}$ is quasi-stationary.

All conditions for Theorem 2B.1 are thus satisfied, and applying this theorem to (8.17) with (8.19) gives the following result.

Lemma 8.2. Consider a uniformly stable, linear model structure $\mathcal{M}$ (see Definitions 4.3 and 8.3). Assume that the data set Z^∞ is subject to D1. Then, with $V_N(\theta, Z^N)$ defined by (8.17),

$$\sup_{\theta \in D_\mathcal{M}} |V_N(\theta, Z^N) - \bar{V}(\theta)| \to 0, \qquad \text{w.p. 1 as } N \to \infty \tag{8.21}$$

where

$$\bar{V}(\theta) = \bar{E}\frac{1}{2}\varepsilon^2(t,\theta) \tag{8.22}$$

■

The criterion function $V_N(\theta, Z^N)$ thus converges uniformly in $\theta \in D_\mathcal{M}$ to the limit function $\bar{V}(\theta)$. This implies that the minimizing argument $\hat{\theta}_N$ of V_N also converges to the minimizing argument θ^* of $\bar{V}$. Notice that it is essential that the convergence is indeed uniform in θ for this to hold (see Problem 8D.1). It may happen that $\bar{V}(\theta)$ does not have a unique global minimum. In that case we define the set of minimizing values as

$$D_c = \arg\min_{\theta \in D_\mathcal{M}} \bar{V}(\theta) = \{\theta | \theta \in D_\mathcal{M}, \bar{V}(\theta) = \min_{\theta' \in D_\mathcal{M}} \bar{V}(\theta')\} \tag{8.23}$$

We can thus formulate this corollary to Lemma 8.2 as our main convergence result:

Theorem 8.2. Let $\hat{\theta}_N$ be defined by (8.16) and (8.17), where $\varepsilon(t,\theta)$ is determined from a uniformly stable linear model structure $\mathcal{M}$. Assume that the data set Z^∞ is subject to D1. Then

$$\hat{\theta}_N \to D_c, \qquad \text{w.p. 1 as } N \to \infty \tag{8.24}$$

where D_c is given by (8.22) and (8.23).

Remark. Convergence into a set as in (8.24) is to be interpreted as

$$\inf_{\bar{\theta} \in D_c} |\hat{\theta}_N - \bar{\theta}| \to 0, \qquad \text{as } N \to \infty \tag{8.25}$$

■

The function $\bar{V}(\theta)$ will in general depend both on the true system and the input properties. With a quadratic criterion and a linear model structure, it follows from Theorem 2.2 that it depends on the data only via the spectrum matrix $\Phi_z(\omega)$ in (8.13). [Explicit expressions will be given in (8.61) to (8.65).] This has the important consequence that it is only the second-order properties of the data that affect the convergence of the estimates.

Ensemble- and Time-averages

The signal sources for $\varepsilon(t,\theta)$ are w and e_0, as evidenced by (8.19). Recall that $w = u$ in case of open loop operation. The symbol $\bar{E}$ denotes as defined in (2.60) ensemble-averaging ("statistical expectation") over the stochastic process $\{e_0(t)\}$ and time-averaging over the deterministic signal $\{w(t)\}$. The function $\bar{V}(\theta)$ is thus "the average value" of $\varepsilon^2(t,\theta)$ in these two respects.

The reason for time-averaging over $\{w(t)\}$ is, as we have stated several times, that it might not always be suitable to describe this signal as a realization of a stochastic process. However, when indeed $\{w(t)\}$ is taken as a realization of a stationary stochastic process, (independent of e_0), Theorem 2.3 shows that, under weak conditions, time averages over $\{w(t)\}$ will, with probability 1, equal the ensemble averages:

$$\lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N} w(t)w(t-\tau) = E_w\, w(t)w(t-\tau) \text{ w.p. } 1 \tag{8.26}$$

Here E_w denotes statistical expectation with respect to the w-process.

This means that e_0-ensemble- and w-time-averaging by $\bar{E}$ will, w.p. 1, be equivalent to taking total statistical expectation over both e_0 and w:

$$\text{"}\bar{E} = E_w E_{e_0} \qquad \text{w.p. 1"} \tag{8.27a}$$

i.e.,

$$\bar{V}(\theta) = \bar{E}\varepsilon^2(t,\theta) = E_w E_{e_0}\varepsilon^2(t,\theta) \qquad \text{w.p. 1} \tag{8.27b}$$

For "hand calculation" it is often easier to apply this total expectation: See Examples 8.1 to 8.2.

[Conversely, one could also replace ensemble averages over e_0 by time averages to eliminate the probabilistic framework entirely: See Problem 8T.1.]

The General Case

With a little more technical effort, the results of Lemma 8.2 and Theorem 8.2 can also be established for general norms $\ell(\varepsilon,\theta)$ as in (7.16), in which case the limit is defined as

$$\bar{V}(\theta) = \bar{E}\ell(\varepsilon(t,\theta),\theta) \tag{8.28}$$

The result can also be extended to nonlinear, time-varying models and less restrictive assumptions on the data set than D1. See, for example, Ljung (1978a) for such results. In summary we thus have

$$\hat{\theta}_N \to \arg\min_{\theta\in D_{\mathcal{M}}} \bar{E}\,\ell(\varepsilon(t,\theta),\theta), \qquad \text{w.p. 1 as } N\to\infty \tag{8.29}$$

This convergence result is quite general and intuitively appealing. It states that the estimate will converge to the *best possible approximation of the system that is available in the model set.* The goodness of the approximation is then measured in terms of the criterion $\bar{V}(\theta)$ in (8.28). We shall dwell on what "best possible" actually means in more practical terms in the next two sections. First we give two examples.

Example 8.1

Suppose that the system is given by

$$y(t) + a_0 y(t-1) = b_0 u(t-1) + e_0(t) + c_0 e_0(t-1) \tag{8.30}$$

where $\{u(t)\}$ and $\{e_0(t)\}$ are independent white noises with unit variances. Let the model structure be given by

$$\hat{y}(t|\theta) + ay(t-1) = bu(t-1), \qquad \theta = \begin{bmatrix} a \\ b \end{bmatrix} \tag{8.31}$$

The prediction-error variance is

$$\begin{aligned} \bar{V}(\theta) &= \bar{E}[y(t) + ay(t-1) - bu(t-1)]^2 \\ &= r_0(1 + a^2 - 2aa_0) + b^2 - 2bb_0 + 2ac_0 \end{aligned} \tag{8.32}$$

where

$$r_0 = Ey^2(t) = \frac{b_0^2 + c_0(c_0 - a_0) - a_0 c_0 + 1}{1 - a_0^2}$$

(see Problem 2E.7). It is easy to verify that the values of a and b that minimize (8.32) are $\theta^* = [a^* \; b^*]^T$ given by

$$\begin{aligned} a^* &= a_0 - \frac{c_0}{r_0} \\ b^* &= b_0 \end{aligned} \tag{8.33}$$

These values give a prediction-error variance

$$\bar{V}(\theta^*) = 1 + c_0^2 - \frac{c_0^2}{r_0} \tag{8.34}$$

This variance is smaller than the "true values" $\theta_0 = [a_0 \; b_0]^T$ inserted into (8.32) would give:

$$\bar{V}(\theta_0) = 1 + c_0^2 \tag{8.35}$$

When we apply the prediction error method to (8.30) and (8.31), the estimates $\hat{a}_N$ and $\hat{b}_N$ will converge, according to Theorem 8.2, to the values given by (8.33). The fact that $a^* \neq a_0$ is usually expressed as that the estimate is "biased." However, it is clear from (8.34) and (8.35) that the bias is beneficial for the prediction performance of the model (8.31). It gives a strictly better predictor for $\hat{a} = a^*$ than for $\hat{a} = a_0$. ■

Example 8.1 stresses that the algorithm indeed gives us the best possible predictor, and it uses its parameters as vehicles for that. It is, however, important to keep in mind that what is the best approximate description of a system in general depends on the input used. We illustrate this by a simple example.

Example 8.2

Consider the system

$$y(t) = b_0 u(t-1) + e_0(t) \tag{8.36}$$

where

$$u(t) = d_0 u(t-1) + w(t) \tag{8.37}$$

and where $\{e_0(t)\}$ and $\{w(t)\}$ are independent white-noise sequences with unit variances. Let the model structure be given by

$$\hat{y}(t|\theta) = bu(t-2), \qquad \theta = b \tag{8.38}$$

The prediction-error variance associated with (8.38) is

$$\begin{aligned} E[y(t) - bu(t-2)]^2 &= E[b_0 u(t-1) - bu(t-2)]^2 + Ee_0^2(t) \\ &= E[(b_0 d_0 - b)u(t-2) + b_0 w(t-1)]^2 + 1 \\ &= \frac{(b_0 d_0 - b)^2}{1 - d_0^2} + b_0^2 + 1 \end{aligned}$$

Hence

$$\hat{b}_N \to b_0 d_0, \qquad \text{w.p. 1 as } N \to \infty$$

since this gives the smallest prediction-error variance. Now the predictor

$$\hat{y}(t|t-1) = b_0 d_0 u(t-2) \tag{8.39}$$

is a fairly reasonable one for the system (8.36) under the input (8.37). It yields the prediction-error variance $1 + b_0^2$, compared to the optimal 1 for a correct model and the output variance

$$1 + \frac{b_0^2}{1 - d_0^2}$$

Notice, however, that the identified model is heavily dependent on the input that was used during the identification experiment. If (8.39) is applied to a white-noise input $\{u(t)\}$, $(d_0 = 0)$, the model (8.39) is useless: It yields the prediction-error variance $1 + b_0^2 + b_0^2 d_0^2$, which is larger than the output variance $1 + b_0^2$. ■

8.4 CONSISTENCY AND IDENTIFIABILITY

Suppose now that assumption S1 holds so that we have a true system, denoted by $\mathcal{S}$. Let us discuss under what conditions it will be possible to recover this system using prediction-error identification.

Clearly, a first assumption must be that $\mathcal{S} \in \mathcal{M}$; that is, the set $D_T(\mathcal{S}, \mathcal{M})$ defined by (8.9) is nonempty.

$\mathcal{S} \in \mathcal{M}$: Quadratic Criteria

The basic consistency result is almost immediate.

Theorem 8.3. Suppose that the data set Z^∞ is subject to assumptions D1 and S1. Let $\mathcal{M}$ be a linear, uniformly stable model structure such that $\mathcal{S} \in \mathcal{M}$. Assume also that Z^∞ is informative enough with respect to $\mathcal{M}$. Then

$$D_c = D_T(\mathcal{S}, \mathcal{M}) \tag{8.40}$$

where D_c is defined by (8.22) and (8.23) and $D_T(\mathcal{S}, \mathcal{M})$ by (8.9). If, in addition, the model structure is globally identifiable at $\theta_0 \in D_T(\mathcal{S}, \mathcal{M})$, then

$$D_c = \{\theta_0\} \tag{8.41}$$ ■

Theorems 8.2 and 8.3 together consequently state that the estimated transfer functions obey

$$G(e^{i\omega},\hat{\theta}_N)\to G_0(e^{i\omega});\ H(e^{i\omega},\hat{\theta}_N)\to H_0(e^{i\omega}),\qquad \text{w.p. 1 as } N\to\infty \tag{8.42}$$

Proof of Theorem 8.3. Let $\theta_0 \in D_T$ and consider, for any $\theta \in D_{\mathcal{M}}$,

$$\bar{V}(\theta) - \bar{V}(\theta_0) = \bar{E}[\varepsilon(t,\theta) - \varepsilon(t,\theta_0)]\varepsilon(t,\theta_0) + \tfrac{1}{2}\bar{E}[\varepsilon(t,\theta) - \varepsilon(t,\theta_0)]^2 \tag{8.43}$$

Since $\theta_0 \in D_T$,

$$\varepsilon(t,\theta_0) = H_0^{-1}(q)G_0(q)u(t) - H_0^{-1}(q)y(t) = e_0(t)$$

according to S1. Moreover, the difference

$$\varepsilon(t,\theta) - \varepsilon(t,\theta_0) = \hat{y}(t|\theta_0) - \hat{y}(t|\theta)$$

depends only on input–output data up to time $t-1$ and is therefore independent of $e_0(t)$ [cf. (8.2)]. The first term of (8.43) is therefore zero. The second term, which equals

$$\bar{E}[\hat{y}(t|\theta_0) - \hat{y}(t|\theta)]^2$$

is strictly positive if θ and θ_0 correspond to different models, since the data set is sufficiently informative; see (8.12). Hence (8.40) follows from (8.23). The result (8.41) follows since global identifiability of $\mathcal{M}$ at θ_0 implies that $D_T = \{\theta_0\}$ [see (4.132)]. ■

$G_0 \in \mathcal{G}$: Quadratic Criteria

Often it is more important to have a good estimate of the transfer function G than of the noise filter H. We shall now study the situation where the set of model transfer functions

$$\mathcal{G} = \{G(e^{i\omega},\theta)|\theta \in D_{\mathcal{M}}\}$$

is large enough to contain the true transfer function,

$$G_0 \in \mathcal{G} \tag{8.44}$$

but the true noise description H_0 cannot be exactly described within the model set. Hence $\mathcal{S} \notin \mathcal{M}$. We then have the following result:

Theorem 8.4. Suppose that the data set Z^∞ is subject to assumptions D1 and S1. Let $\mathcal{M}$ be a linear uniformly stable model structure, such that G and H are independently parametrized:

$$\theta = \begin{bmatrix}\rho\\ \eta\end{bmatrix}\qquad G(q,\theta) = G(q,\rho),\qquad H(q,\theta) = H(q,\eta) \tag{8.45}$$

and such that the set

$$D_G(\mathcal{S},\mathcal{M}) = \{\rho\,|\,G(e^{i\omega},\rho) = G_0(e^{i\omega})\ \forall\omega\} \tag{8.46}$$

is nonempty. Assume that Z^∞ is informative enough with respect to $\mathcal{M}$ and that the system operates in open loop; that is,

$$\{u(t)\} \text{ and } \{e_0(t)\} \text{ are independent} \tag{8.47}$$

Let

$$\hat{\theta}_N = \begin{bmatrix} \hat{\rho}_N \\ \hat{\eta}_N \end{bmatrix}$$

be obtained by the prediction-error method (8.16) and (8.17). Then

$$\hat{\rho}_N \to D_G(\mathcal{S}, \mathcal{M}), \qquad \text{w.p. 1 as } N \to \infty \tag{8.48}$$

■

The result (8.48) can be written more suggestively as

$$G(e^{i\omega}, \hat{\theta}_N) \to G_0(e^{i\omega}), \qquad \text{w.p. 1 as } N \to \infty \tag{8.49}$$

Proof. Consider the function $\overline{V}(\theta)$ given by (8.22). We have from S1

$$\begin{aligned} \varepsilon(t,\theta) &= H^{-1}(q,\eta)[y(t) - G(q,\rho)u(t)] \\ &= H^{-1}(q,\eta)[(G_0(q) - G(q,\rho))u(t) + H_0(q)e_0(t)] \\ &= u_F(t,\eta,\rho) + e_F(t,\eta) \end{aligned}$$

with obvious notation.

Since u and e_0 are independent, we have that

$$\overline{V}(\theta) = \overline{V}(\rho,\eta) = \tfrac{1}{2}[Eu_F^2(t,\rho,\eta) + Ee_F^2(t,\eta)]$$

The first term is zero precisely when $\rho \in D_G(\mathcal{S}, \mathcal{M})$, and the second term is independent of ρ. Hence

$$\arg\min_{\rho} \overline{V}(\rho,\eta) = D_G(\mathcal{S}, \mathcal{M})$$

irrespective of H, which, together with Theorem 8.2, concludes the proof. ■

We may add that both assumptions (8.45) and (8.47) are essential for the result to hold. See Example 8.1 and Problem 8E.3.

The case of independent parametrization (8.45) covers the output error model (4.25) along with variants with fixed noise models

$$y(t) = G(q,\theta)u(t) + H_*(q)e(t) \tag{8.50}$$

[which alternatively can be regarded as the output error model used with a prefilter $L(q) = 1/H_*(q)$; see (7.13) and (7.14)]. It also covers the Box–Jenkins model structure (4.31). These model structures consequently have the important advantage that the transfer function G can be consistently estimated, even when the noise model set is too simple to admit a completely correct description of the system.

Example 8.3

Consider the system (8.30) of Example 8.1, and let the model structure be a first-order output error model:

$$\hat{y}(t|\theta) = \frac{bq^{-1}}{1 + aq^{-1}} u(t)$$

In this case it follows from Theorem 8.4 that the estimates $\hat{a}_N$ and $\hat{b}_N$ will converge to the true values a_0 and b_0. ■

$\mathcal{S} \in \mathcal{M}$: General Norm $\ell(\varepsilon)$ (*)

With a general, θ-independent norm $\ell(\varepsilon)$, the estimate converges into the set D_c:

$$\hat{\theta}_N \to D_c = \arg\min_{\theta \in D_{\mathcal{M}}} \bar{E}\ell(\varepsilon(t, \theta)) \tag{8.51}$$

according to (8.29). In general, the set D_c will depend on ℓ. However, when $\mathcal{S} \in \mathcal{M}$ it is desirable that $D_c = D_T(\mathcal{S}, \mathcal{M})$ for all reasonable choices of ℓ. Clearly, some conditions must be imposed on ℓ, and Problem 8D.3 shows that it is not sufficient to require $\ell(\varepsilon)$ to be increasing with $|\varepsilon|$. We have to require $\ell(\varepsilon)$ to be convex to be able to prove a result that holds for all distributions of the innovation $e_0(t)$. We thus have the following extension of Theorem 8.3.

Theorem 8.5. Let $\ell(x)$ be a twice differentiable function such that

$$E\ell'(e_0(t)) = 0 \tag{8.52}$$

$$\ell''(x) \geq \delta > 0, \qquad \forall x \tag{8.53}$$

Here $e_0(t)$ are the innovations in assumption S1. Then, under the assumptions of Theorem 8.3,

$$D_c = D_T(\mathcal{S}, \mathcal{M})$$

with D_c defined by (8.28) and (8.23).

Proof. Let $\theta_0 \in D_T$ and denote as usual

$$\varepsilon(t, \theta_0) = e_0(t)$$

Then for any $\theta \notin D_T$

$$\varepsilon(t, \theta) = e_0(t) + \tilde{y}(t, \theta)$$

where $\bar{E}[\tilde{y}(t, \theta)]^2 > 0$ since the data set is sufficiently informative. Hence, by Taylor's expansion,

$$\ell(\varepsilon(t, \theta)) = \ell(e_0(t)) + \tilde{y}(t, \theta)\ell'(e_0(t)) + \tfrac{1}{2}[\tilde{y}(t, \theta)]^2 \ell''(\xi(t))$$

where $\xi(t)$ is a value between $e_0(t)$ and $\varepsilon(t, \theta)$. Since $e_0(t)$ and $\tilde{y}(t, \theta)$ are independent, this expression gives

$$\begin{aligned}\overline{E}\ell(\varepsilon(t, \theta)) &= \overline{E}\ell(e_0(t)) + 0 + \tfrac{1}{2}\overline{E}\{[\tilde{y}(t, \theta)]^2 \ell''(\xi(t))\} \\ &\geq \overline{E}\ell(e_0(t)) + \delta \cdot \overline{E}[\tilde{y}(t, \theta)]^2 > \overline{E}\ell(e_0(t))\end{aligned}$$

using (8.52) and (8.53) and $\overline{E}[\tilde{y}(t, \theta)]^2 > 0$, respectively. This concludes the proof. ■

Clearly, an analogous extension of Theorem 8.4 can also be given.

In the maximum likelihood method the norm ℓ is chosen as the negative logarithm of the PDF of the innovations; (7.73):

$$\ell(x) = -\log f_e(x) \tag{8.54}$$

It can be shown that (8.52) automatically holds for this norm, and that Theorem 8.5 holds without condition (8.53). See Problem 8G.3.

$\mathcal{S} \in \mathcal{M}$: General Norm $\ell(\varepsilon, \alpha)$ (*)

We consider now the case where the norm is parametrized by an α that is independent of the parametrization of the predictor as in (7.17). We this have that the limit values of θ and α are given by

$$(\theta^*, \alpha^*) = \arg\min_{\theta, \alpha} \overline{V}(\theta, \alpha) = \arg\min_{\theta, \alpha} \overline{E}\ell(\varepsilon(t, \theta), \alpha) \tag{8.55}$$

If $\mathcal{S} \in \mathcal{M}$ and the conditions of Theorem 8.5 are satisfied for all α, then it is clear that $\theta^* \in D_T(\mathcal{S}, \mathcal{M})$, regardless of α. This means that

$$\alpha^* = \arg\min_{\alpha} \overline{E}\ell(e_0(t), \alpha) \tag{8.56}$$

We shall study what (8.56) tells us about the limit value α^*. We first have the following result.

Lemma 8.3. Consider a norm (7.17), normalized so that

$$\int_{-\infty}^{\infty} e^{-\ell(x, \alpha)}\, dx = 1 \ \forall \alpha \tag{8.57}$$

Let the PDF of $e_0(t)$ be $f_e(x)$, and assume that for some α_0

$$\ell(x, \alpha_0) = -\log f_e(x) \tag{8.58}$$

Then $\alpha^* = \alpha_0$ in (8.56).

Proof. Let

$$f_\alpha(x) = e^{-\ell(x, \alpha)}$$

Hence

$$\bar{E}\ell(e_0(t),\alpha) - \bar{E}\ell(e_0(t),\alpha_0) = -\bar{E}\log\frac{f_\alpha(e_0(t))}{f_e(e_0(t))}$$

$$\geq -\log E\frac{f_\alpha(e_0(t))}{f_e(e_0(t))} = -\log\int\left[\frac{f_\alpha(x)}{f_e(x)}\right]f_e(x)dx = -\log\int f_\alpha(x)dx = 0$$

The inequality is Jensen's inequality (see Chung, 1974) since $\log x$ is a convex function, and equality holds if and only if $f_\alpha(x) = \text{const}\cdot f_e(x)$. This proves the lemma. ■

Heuristically, we could thus say that

the minimization with respect to α in (8.55) tries to make the norm $\ell(\varepsilon,\alpha)$ look like the negative logarithm of the PDF of the true innovations. (8.59)

Example 8.4

Let $\ell(\varepsilon,\alpha)$ be given by (7.75):

$$\ell(\varepsilon,\alpha) = \frac{1}{2}\left[\frac{\varepsilon^2}{\alpha} + \log\alpha\right]$$

We find that

$$\bar{E}[\ell(e_0(t),\alpha)] = \frac{1}{2}\left[\frac{Ee_0^2(t)}{\alpha} + \log\alpha\right] = \frac{1}{2}\left[\frac{\lambda_0}{\alpha} + \log\alpha\right]$$

which is minimized by $\alpha = \lambda_0$. The estimate $\hat{\alpha}_N$ will thus converge to the innovation variance as N tends to infinity. See also Problem 7E.7. ■

When the parametrizations of the predictor and of the norm $\ell(\varepsilon,\theta)$ have common parameters, the conclusion is that

$$\theta^* = \arg\min_{\theta\in D_{\mathcal{M}}} \bar{E}\ell(\varepsilon(t,\theta),\theta)$$

will give a compromise between making the prediction errors $\{\varepsilon(t,\theta)\}$ equal to the true innovations $\{e_0(t)\}$ and (8.59), that is, making the norm look like $-\log f_e(x)$. In case these two objections cannot be reached simultaneously, consistency may be lost even if $D_T(\mathcal{S},\mathcal{M})$ is nonempty. See Problem 8E.2.

Multivariable Case (*)

The convergence and consistency results for multivariable systems are entirely analogous to the scalar case. The result (8.29) holds without notational changes for the multivariable case. The counterparts of Theorems 8.3 and 8.4 with quadratic criteria

$$\bar{E}\ell(\varepsilon(t,\theta)) = \tfrac{1}{2}\bar{E}\varepsilon^T(t,\theta)\Lambda^{-1}\varepsilon(t,\theta) \tag{8.60}$$

hold as stated, with only obvious notational changes in the proofs. For Theorem 8.5, the condition (8.53) takes the form that the $p \times p$ matrix $\ell''(\varepsilon)$ should be positive definite.

8.5 LINEAR TIME-INVARIANT MODELS: A FREQUENCY-DOMAIN DESCRIPTION OF THE LIMIT MODEL

Theorem 8.2 describes the limiting estimate θ^*, $\theta^* \in D_c$, as the one that minimizes the prediction error variance among all models in the structure $\mathcal{M}$. In case $\mathcal{S} \in \mathcal{M}$, this means that $\theta^* = \theta_0$ is a true description of the system (see Theorem 8.3), but otherwise the model will differ from the true system. In this section we shall develop some expressions that characterize this misfit between the limiting model and the true system for the case of linear time-invariant models. See also Problem 8G.4.

The Function $\bar{V}(\theta)$ for Open-Loop Operation

By the fundamental expression (2.65), we may write

$$\bar{V}(\theta) = \bar{E}\frac{1}{2}\varepsilon^2(t,\theta) = \frac{1}{4\pi}\int_{-\pi}^{\pi} \Phi_\varepsilon(\omega,\theta)d\omega \tag{8.61}$$

where $\Phi_\varepsilon(\omega,\theta)$ is the spectrum of the prediction errors $\{\varepsilon(t,\theta)\}$. Under assumption S1 we have

$$y(t) = G_0(q)u(t) + v_0(t) \tag{8.62}$$

where the additive noise $\{v_0(t)\}$ has the spectrum

$$\Phi_v(\omega) = \lambda_0 |H_0(e^{i\omega})|^2 \tag{8.63}$$

Then, for a linear model structure, we obtain the prediction errors

$$\begin{aligned}\varepsilon(t,\theta) &= H^{-1}(q,\theta)[y(t) - G(q,\theta)u(t)] \\ &= H^{-1}(q,\theta)[(G_0(q) - G(q,\theta))u(t) + v_0(t)]\end{aligned} \tag{8.64}$$

Applying Theorem 2.2 to (8.64) gives the expression for $\Phi_\varepsilon(\omega,\theta)$:

$$\Phi_\varepsilon(\omega,\theta) = \frac{|\tilde{G}(e^{i\omega},\theta)|^2 \Phi_u(\omega) + \Phi_v(\omega)}{|H(e^{i\omega},\theta)|^2} \tag{8.65}$$

provided v_0 and u are independent. Here $\tilde{G}$ is the difference in (8.15). Hence

$$\begin{aligned}\bar{V}(\theta) &= \bar{E}\frac{1}{2}\varepsilon^2(t,\theta) \\ &= \frac{1}{4\pi}\int_{-\pi}^{\pi} [|G_0(e^{i\omega}) - G(e^{i\omega},\theta)|^2 \Phi_u(\omega) + \Phi_v(\omega)]\frac{1}{|H(e^{i\omega},\theta)|^2}d\omega\end{aligned} \tag{8.66}$$

We now have a characterization of

$$D_c = \arg\min_{\theta} \bar{V}(\theta) \tag{8.67}$$

in the frequency domain.

Fixed Noise Model

For a fixed noise model $H(q,\theta) = H^*(q)\ \forall\theta$, (8.66) and (8.67) can be rewritten:

$$D_c = \arg\min_{\theta} \int_{-\pi}^{\pi} |G_0(e^{i\omega}) - G(e^{i\omega},\theta)|^2 Q^*(\omega) d\omega \tag{8.68a}$$

$$Q^*(\omega) = \frac{\Phi_u(\omega)}{|H^*(e^{i\omega})|^2} \tag{8.68b}$$

where we disposed of θ-independent terms. Let $\theta^* \in D_c$. In this case, the limiting model

$$G^*(e^{i\omega}) = G(e^{i\omega},\theta^*)$$

is a clear-cut best mean-square approximation of $G_0(e^{i\omega})$, with a frequency weighting Q^* that depends on the noise model H^* and the input spectrum Φ_u, and can be interpreted as the model signal-to-noise ratio.

Independently Parametrized Noise Model (*)

Consider now an independently parametrized noise model (8.45), (4.125). Let $\theta^* \in D_c$. Then we can write $\theta^* = [\rho^*\ \eta^*]^T$.

$$\rho^* = \arg\min_{\rho} \int_{-\pi}^{\pi} |G_0(e^{i\omega}) - G(e^{i\omega},\rho)|^2 Q(\omega,\eta^*) d\omega \tag{8.69a}$$

$$Q(\omega,\eta) = \frac{\Phi_u(\omega)}{|H(e^{i\omega},\eta)|^2} \tag{8.69b}$$

$$\eta^* = \arg\min_{\eta} \int_{-\pi}^{\pi} \frac{\Phi_{ER}(\omega,\rho^*)}{|H(e^{i\omega},\eta)|^2} d\omega \tag{8.70a}$$

where

$$\Phi_{ER}(\omega,\rho) = |G_0(e^{i\omega}) - G(e^{i\omega},\rho)|^2 \Phi_u(\omega) + \Phi_v(\omega) \tag{8.70b}$$

is the *error spectrum*, i.e., the spectrum of the output error $y(t) - G(q,\rho)u(t)$. In this case we see that $G^*(e^{i\omega})$ is fitted to $G_0(e^{i\omega})$ in the $Q(\omega,\eta^*)$ norm, which is not known a priori, but defined indirectly through (8.70) via the noise model $H(e^{i\omega},\eta^*)$.

To better understand the minimization problem (8.70), let us factorize the error spectrum

$$\Phi_{ER}(\omega,\rho^*) = \lambda^* |N(e^{i\omega},\rho^*)|^2 \tag{8.71}$$

for a monic, stable and inversely stable $N(q, \rho^*)$. Notice that when $G(q, \rho^*) = G_0(q)$ then $N(q, \rho^*)$ equals $H_0(q)$; see (8.70b) and (8.63). Then

$$\begin{aligned}\frac{\Phi_{ER}(\omega,\rho^*)}{|H(e^{i\omega},\eta)|^2} &= \lambda^* \left| \frac{N(e^{i\omega},\rho^*)}{H(e^{i\omega},\eta)} \right|^2 \\ &= \lambda^* \left| 1 + \frac{N(e^{i\omega},\rho^*) - H(e^{i\omega},\eta)}{H(e^{i\omega},\eta)} \right|^2 \\ &= \lambda^* [1 + |R(e^{i\omega})|^2 + R(e^{i\omega}) + \overline{R(e^{i\omega})}]\end{aligned} \tag{8.72}$$

with

$$R(e^{i\omega}) = \frac{N(e^{i\omega},\rho^*) - H(e^{i\omega},\eta)}{H(e^{i\omega},\eta)} \tag{8.73}$$

(R depends on η and ρ^*, but we drop these arguments). Since N and H are both monic and since $H(e^{i\omega},\eta)$ is inversely stable, we can write

$$R(e^{i\omega}) = \sum_{k=1}^{\infty} r(k) e^{-ik\omega}$$

the important observation being that the term corresponding to $k = 0$ in the sum is zero. Thus

$$\int_{-\pi}^{\pi} R(e^{i\omega}) d\omega = 0$$

Note also that we may write

$$R(e^{i\omega}) = -\left[\frac{1}{N(e^{i\omega},\rho^*)} - \frac{1}{H(e^{i\omega},\eta)} \right] \cdot N(e^{i\omega},\rho^*) \tag{8.74}$$

Collecting all this means that (8.70) can be rewritten

$$\eta^* = \arg\min_{\eta} \int_{-\pi}^{\pi} \left| \frac{1}{N(e^{i\omega},\rho^*)} - \frac{1}{H(e^{i\omega},\eta)} \right|^2 \cdot \Phi_{ER}(\omega,\rho^*) d\omega \tag{8.75}$$

which shows that the inverse noise model $1/H$ is fitted to be the inverse spectral factor of the error spectrum $1/N$ (defined by (8.71)) in a quadratic norm given by the error spectrum. Put less formally,

> η^* is such that the model noise spectrum $|H(e^{i\omega},\eta^*)|^2$ resembles the error spectrum $\Phi_{ER}(\omega,\rho^*)$ as much as possible, within the chosen set of model noise spectra. (8.76)

General Case

In the general case (8.66), when the noise model has parameters in common with the transfer function, no clear-cut formal characterization of the resulting estimates can be given. It is useful, though, and intuitively appealing to see the

resulting estimate θ^* as a compromise between fitting $G(e^{i\omega}, \theta)$ to $G_0(e^{i\omega})$ in the quadratic frequency norm,

$$Q(\omega, \theta^*) = \frac{\Phi_u(\omega)}{|H(e^{i\omega}, \theta^*)|^2} \tag{8.77}$$

and fitting the model spectrum $|H(e^{i\omega}, \theta)|^2$ to the error spectrum,

$$\Phi_{ER}(\omega, \theta^*) = |G_0(e^{i\omega}) - G(e^{i\omega}, \theta^*)|^2 \Phi_u(\omega) + \Phi_v(\omega) \tag{8.78}$$

in the manner described by (8.75). This interpretation, although approximate, is quite illuminating.

Example 8.5

Consider the system

$$y(t) = G_0(q)u(t)$$

with

$$G_0(q) = \frac{0.001q^{-2}(10 + 7.4q^{-1} + 0.924q^{-2} + 0.1764q^{-3})}{1 - 2.14q^{-1} + 1.553q^{-2} - 0.4387q^{-3} + 0.042q^{-4}} \tag{8.79}$$

No disturbances act on the system. The input is a PRBS (see Chapter 14) with basic period one sample, which gives $\Phi_u(\omega) \approx 1$ all ω.

This system was identified with the prediction-error method using a quadratic criterion and prefilter $L(q) \equiv 1$ in the output error model structure

$$\hat{y}(t|\theta) = \frac{b_1 q^{-1} + b_2 q^{-2}}{1 + f_1 q^{-1} + f_2 q^{-2}} u(t) \tag{8.80}$$

Bode plots of the true system and of the resulting model are given in Figure 8.2. We see that the model gives a good description of the low-frequency properties but is bad at high frequencies. According to (8.68), the limiting model is characterized by

$$\theta^* = \arg \min_{\theta} \int_{-\pi}^{\pi} |G_0(e^{i\omega}) - G(e^{i\omega}, \theta)|^2 \, d\omega \tag{8.81}$$

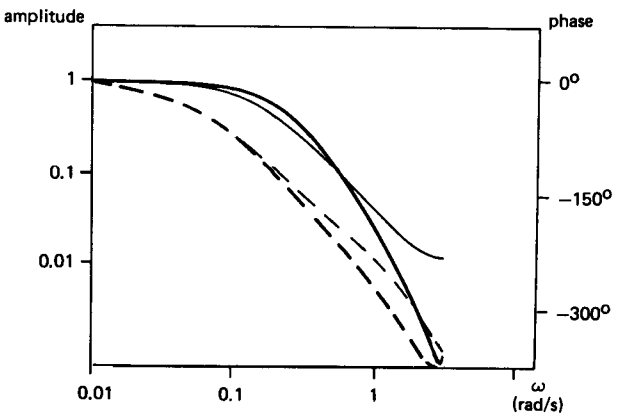

Figure 8.2 Bode plots of true system and model identified in (8.80). Solid lines: Amplitude plots; dashed lines: phase plots; thick lines: true system; thin lines: estimate.

since $H^*(q) = 1$ and $\Phi_u(\omega) \equiv 1$. Since the amplitude of the true system falls off by a factor of 10^{-2} to 10^{-3} for $\omega > 1$, it is clear that errors at higher frequencies contribute only marginally to the criterion (8.81); hence the good low-frequency fit.

Consider now instead an ARX structure

$$y(t) = \frac{b_1 q^{-1} + b_2 q^{-2}}{1 + a_1 q^{-1} + a_2 q^{-2}} u(t) + \frac{1}{1 + a_1 q^{-1} + a_2 q^{-2}} e(t) \tag{8.82}$$

corresponding to the linear regression predictor

$$\hat{y}(t|\theta) = -a_1 y(t-1) - a_2 y(t-2) + b_1 u(t-1) + b_2 u(t-2)$$

When applied to the same data, this structure gives the model description in Figure 8.3, with a much worse low-frequency fit. According to our discussion in this section, this limit model is a compromise between fitting $1/|1 + a_1 e^{i\omega} + a_2 e^{2i\omega}|^2$ to the error spectrum and minimizing (a_1^* and a_2^* correspond to the limit estimate θ^*)

$$\int_{-\pi}^{\pi} |G_0(e^{i\omega}) - G(e^{i\omega}, \theta)|^2 \cdot |1 + a_1^* e^{i\omega} + a_2^* e^{2i\omega}|^2 \, d\omega \tag{8.83}$$

The function $|A_*(e^{i\omega})|^2 = |1 + a_1^* e^{i\omega} + a_2^* e^{2i\omega}|^2$ is plotted in Figure 8.4. It assumes values at high frequencies that are 10^4 times those at low frequencies. Hence, compared to (8.81), the

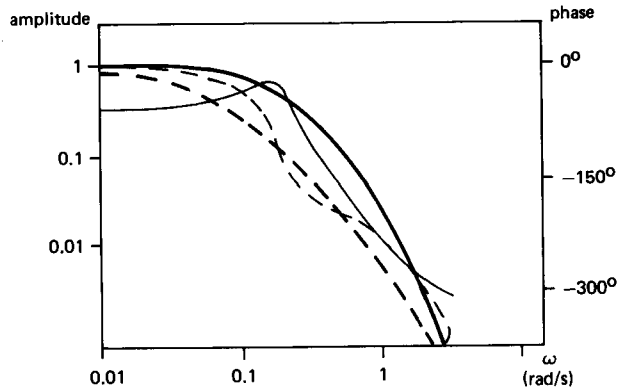

Figure 8.3 Bode plots of true system and of model identified in (8.82). Legend as in Figure 8.2.

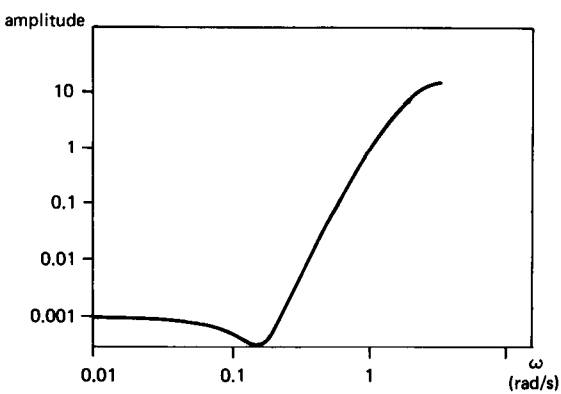

Figure 8.4 The weighting function $|A_*(e^{i\omega})|^2$ in (8.83).

criterion (8.83) penalizes high-frequency misfit much more. This explains the different properties of the limit models obtained in the model structures (8.80) and (8.82), respectively. ■

8.6 THE CORRELATION APPROACH

In Section 7.5 we defined the correlation approach to identification, with the special cases of PLR and IV methods. The convergence analysis for these methods is quite analogous to the prediction-error approach as given in the previous few sections.

Basic Convergence Result

Consider the function

$$f_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} \zeta(t, \theta)\varepsilon_F(t, \theta) \tag{8.84}$$

where ε_F is given by

$$\varepsilon_F(t, \theta) = L(q)\varepsilon(t, \theta) \tag{8.85}$$

and the correlation vector $\zeta(t, \theta)$ is obtained by linear filtering of past data:

$$\zeta(t, \theta) = K_y(q, \theta)y(t) + K_u(q, \theta)u(t) \tag{8.86}$$

(both filters contain one delay). Determining the estimate $\hat{\theta}_N$ by solving $f_N(\theta, Z^N) = 0$ gives the correlation approach (7.96). We have here specialized the general instruments (7.96b) to a linear case (8.86).

The convergence analysis of (8.84) is entirely analogous to the prediction-error case. Thus we have from Theorem 2B.1:

Lemma 8.4. Let the data set Z^∞ be subject to D1, and let the prediction errors be computed using a uniformly stable linear model structure. Assume that the family of filters

$$\{K_y(q, \theta), K_u(q, \theta);\ \theta \in D_{\mathcal{M}}\}$$

is uniformly stable. Then

$$\sup_{\theta \in D_{\mathcal{M}}} |f_N(\theta, Z^N) - \bar{f}(\theta)| \to 0, \qquad \text{w.p. 1 as } N \to \infty \tag{8.87}$$

where

$$\bar{f}(\theta) = \bar{E}\zeta(t, \theta)\varepsilon_F(t, \theta) \qquad \blacksquare \tag{8.88}$$

For the estimate $\hat{\theta}_N$, we thus have the following result.

Theorem 8.6. Let $\hat{\theta}_N$ be defined by

$$\hat{\theta}_N = \operatorname*{sol}_{\theta \in D_{\mathcal{M}}} [f_N(\theta, Z^N) = 0]$$

Then, subject to the assumptions of Lemma 8.4,

$$\hat{\theta}_N \to D_{cf}, \qquad \text{w.p. 1 as } N \to \infty \tag{8.89}$$

where

$$D_{cf} = \{\theta | \theta \in \acute{D}_{\mathcal{M}}, \bar{f}(\theta) = 0\} \qquad \blacksquare \tag{8.90}$$

The theorem is here given for the special choice $\alpha(\varepsilon_F) = \varepsilon_F$ in (7.96d). The extension to general $\alpha(\cdot)$ is straightforward.

This convergence result is quite general and also quite natural. The limiting estimate $\theta^* \in D_{cf}$ will be characterized by the property that the filtered prediction errors $\{\varepsilon_F(t, \theta^*)\}$ indeed are uncorrelated with the instruments $\{\zeta(t, \theta^*)\}$. This was also our guideline when selecting $\hat{\theta}_N$. We shall now characterize D_{cf} in more practical terms for some special cases.

$\mathcal{S} \in \mathcal{M}$: Consistency (*)

An assumption $\mathcal{S} \in \mathcal{M}$ would in this case be that there exists a value $\theta_0 \in D_{\mathcal{M}}$ such that $\{\varepsilon(t, \theta_0) = e_0(t)\}$ is a white noise. With $L(q) = 1$, we thus find that $\bar{f}(\theta_0) = 0$, since $e_0(t)$ is independent of past data and in particular of $\zeta(t, \theta_0)$. Hence, as expected.

$$\theta_0 \in D_{cf} \tag{8.91}$$

Whether this set contains more elements when the data are informative and the model structure globally identifiable at θ_0 is not so easy to analyze in the general case.

$G_0 \in \mathcal{G}$: Instrumental-variable Methods

Consider the IV method with instruments

$$\zeta(t, \theta) = K_u(q, \theta)u(t) \tag{8.92}$$

The underlying model is

$$A(q)y(t) = B(q)u(t) + v(t)$$

for which the predictor

$$\hat{y}(t|\theta) = \varphi^T(t)\theta$$

is determined as in (4.11) and (4.12).

Under assumption S1, the true system is given by (8.7). If there exists a θ_0 corresponding to $(A_0(q), B_0(q))$ such that

$$G_0(q) = \frac{B_0(q)}{A_0(q)}, \qquad [G_0 \in \mathcal{G}; \text{ cf. (8.44)}]$$

we can consequently write (8.7) as

$$y(t) = \frac{B_0(q)}{A_0(q)}u(t) + H_0(q)e_0(t) \tag{8.93}$$

or

$$y(t) = \varphi^T(t)\theta_0 + w_0(t) \tag{8.94}$$

where

$$w_0(t) = A_0(q)H_0(q)e_0(t) \tag{8.95}$$

Suppose now that the system operates in open loop so that $\{w_0(t)\}$ and $\{u(t)\}$ are independent. Then

$$\begin{aligned}\bar{f}(\theta) &= \bar{E}\zeta(t, \theta)L(q)[\varphi^T(t)(\theta_0 - \theta) + w_0(t)] \\ &= [\bar{E}\zeta(t, \theta)\varphi_F^T(t)](\theta_0 - \theta)\end{aligned} \tag{8.96}$$

where

$$\varphi_F^T(t) = L(q)\varphi^T(t) \tag{8.97}$$

The second equality in (8.96) follows since $\zeta(t, \theta)$ is entirely constructed from past $u(t)$, while $L(q)w_0(t)$ is independent of $\{u(t)\}$. Under the stated assumptions, we thus have that $\theta_0 \in D_{cf}$, and whether this set contains more θ-values depends on whether the matrix $\bar{E}\zeta(t, \theta)\varphi_F^T(t)$ is singular.

Suppose now that the instruments ζ do not depend on θ and are generated according to (7.108) to (7.110). The matrix

$$R = \bar{E}\zeta(t)\varphi_F^T(t) \tag{8.98}$$

is then a constant matrix that depends only on the filters $L(q)$, $K(q)$, $N(q)$, and $M(q)$, on the true system, and on the properties of $\{u(t)\}$. A thorough discussion of the nonsingularity if R is given in Söderström and Stoica (1983). We first note the following facts. Let n_a^0, n_b^0 be the orders of the true description (8.93), and let n_a, n_b be the corresponding model orders. Let the orders of the instrument filters (7.108) to (7.110) be n_n and n_m. Then

1. If $\min(n_a - n_a^0, n_b - n_b^0) > 0$, then R is singular. (8.99a)
2. If $\min(n_a - n_n, n_b - n_m) > 0$, then R is singular. (8.99b)

To see this, let

$$z_0(t) = \frac{B_0(q)}{A_0(q)}u(t) \tag{8.100}$$

$$\tilde{\varphi}(t) = [-z_0(t-1) \dots -z_0(t-n_a)\, u(t-1) \dots u(t-n_b)]^T$$

Let $\tilde{\varphi}_F(t) = L(q)\,\tilde{\varphi}(t)$. If $n_a > n_a^0$ and $n_b > n_b^0$, then (8.100) implies that there exists

an $(n_a + n_b)$-dimensional vector S such that $\tilde{\varphi}^T(t)S \equiv 0$. Then also $\tilde{\varphi}_F^T(t)S \equiv 0$. Now, since $\{w_0(t)\}$ and $\{u(t)\}$ are independent, we have

$$R = \bar{E}\zeta(t)\varphi_F^T(t) = \bar{E}\zeta(t)\tilde{\varphi}_F^T(t) \tag{8.101}$$

which shows that $RS = 0$ and R is singular. Similarly, (8.99b) implies the existence of a vector S such that $S^T\zeta(t) \equiv 0$.

When neither of (8.99) hold, the matrix R is "generically" nonsingular. To show this, the reasoning goes as follows: For a given true system and a given input, denote the coefficients of the filters L, K, N, and M by ρ. The matrix R is thus a function of $\rho: R(\rho)$. Now consider the scalar-valued function $\det R(\rho)$. This is an analytic function of ρ (see Finigan and Rowe, 1974). If such a function is zero for ρ in a set of positive Lebesque measure, then it must be identically zero. If we can find a value ρ^* such that $\det R(\rho^*) \neq 0$, we thus can conclude that $\det R(\rho) \neq 0$ for almost all ρ (in the set where $\det R$ is analytic). Such ρ^* can be found if the input spectrum $\Phi_u(\omega) > 0$ for all ω and the orders of the filters N and M are chosen at least as large as the corresponding model orders n_a and n_b (see Problem 8T.2). We thus have the following result.

> Suppose that the system is given by (8.93), that $\Phi_u(\omega) > 0$, and that u and e_0 are independent. Let the instruments $\zeta(t)$ be given by (7.108) to (7.110). Assume that neither of the conditions in (8.99) holds. Then R in (8.98) is nonsingular for almost all such choices of N, M, L, and K. (8.102)

Frequency-domain Characterization of D_{cf} for the IV Method (*)

The prediction errors, under assumption S1, can be written

$$\begin{aligned}\varepsilon(t, \theta) &= A(q)y(t) - B(q)u(t)\\ &= A(q)\left[G_0(q)u(t) + H_0(q)e_0(t) - \frac{B(q)}{A(q)}u(t)\right]\end{aligned}$$

using (8.93). With the instruments given by (8.92), we thus have, analogous to (8.61) to (8.66),

$$\begin{aligned}\bar{f}(\theta) &= \bar{E}\zeta(t, \theta)\varepsilon_F(t, \theta)\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}[G_0(e^{i\omega}) - G(e^{i\omega}, \theta)]\Phi_u(\omega)\cdot A(e^{i\omega})\cdot L(e^{i\omega})\cdot K_u(e^{-i\omega}, \theta)d\omega\end{aligned} \tag{8.103}$$

with

$$G(e^{i\omega}, \theta) = \frac{B(e^{i\omega})}{A(e^{i\omega})}$$

Here $K_u(e^{-i\omega}, \theta)$ is a d-dimensional column vector.

The limiting estimates $\theta_{IV}^* \in D_{cf}$ are thus characterized by the fact that certain scalar products with the error $G_0(e^{i\omega}) - G(e^{i\omega}, \theta_{IV}^*)$ over the frequency domain are zero.

8.7 SUMMARY

In this chapter we have answered the question of what would happen with the estimates if more and more observed data become available. The answer is natural: We have for the prediction-error approach (7.120) that

$$\hat{\theta}_N \to \arg\min_{\theta \in D_M} \bar{E}\ell(\varepsilon(t, \theta), \theta), \qquad \text{w.p. 1 as } N \to \infty \tag{8.104}$$

and for the correlation approach (7.121) that

$$\hat{\theta}_N \to \operatorname{sol}_{\theta \in D_M} [\bar{E}\zeta(t, \theta)\alpha(\varepsilon_F(t, \theta)) = 0], \qquad \text{w.p. 1 as } N \to \infty \tag{8.105}$$

These results were proved in Theorems 8.2 and 8.6, respectively.

The limiting models are thus the same as those that we could have computed as the best system approximations, in case the probabilistic properties of the system had been known to begin with.

In case a true description of the system is available in the model set, the model will converge to this description under certain natural conditions on ℓ, provided the data set is informative enough. This was shown in Theorems 8.3 to 8.5.

When no exact description can be obtained, the model will be fitted to the system in a way that for linear time-invariant models can be characterized in the frequency domain as follows [see (8.68) to (8.76)]:

> The limiting transfer function estimate $G^*(e^{i\omega})$ is partly or entirely determined as the closest function to the true transfer function, measured in a quadratic norm over the frequency range, with a weighting function $\Phi_u(\omega)/|H(e^{i\omega}, \theta^*)|^2$, while the resulting noise model $|H(e^{i\omega}, \theta^*)|^2$ resembles the error spectrum $\Phi_{ER}(\omega)$ as much as possible. (8.106)

8.8 BIBLIOGRAPHY

The convergence and consistency analysis of parameter estimates has, of course, a long history in the statistical literature. The analysis given here has some roots in consistency analysis of the maximum likelihood method, starting with Wald (1949) for the case of independent samples and Åström and Bohlin (1965) for applications to dynamical systems. Analysis of the maximum likelihood method applied to time series is, essentially, equivalent. This problem is described in Walker (1964), Hannan (1970), Anderson (1971), Rissanen and Caines (1979), and many other contributions. Another study that is relevant for our investigation is Jennrich's (1969) analysis of nonlinear least squares. Asymptotic analysis, which does not utilize a stochastic framework, is given in Niederlinski (1984) and Ljung (1985e).

Analysis in the prediction-error case, where the distribution of the disturbances is not necessarily Gaussian, has been given in Hannan (1973), Ljung (1974, 1976c), and Caines (1976a). The situation, studied in the present chapter, where a true description is not necessarily available in the model set, was, perhaps, first analyzed in Ljung (1976a, 1978a). Similar results are also given in Caines (1978), Anderson, Moore, and Hawkes (1978), and Kabaila and Goodwin (1980).

The frequency-domain characterization of the limiting model is further discussed in Wahlberg and Ljung (1986).

Convergence and consistency of IV estimates are treated in a very comprehensive manner in Söderström and Stoica (1983). A variant of the "generic" consistency result for IV estimates (8.102) was first given by Finigan and Rowe (1974). The first consistency analysis of the IV method when applied to dynamic systems was apparently given by Wong and Polak (1967).

It could be noted that, when the prediction-error method is applied to linear regression so that the least-squares method results, the analysis simplifies considerably. The first analysis of this situation was by Mann and Wald (1943). Recent contributions on the consistency of the LS method include Ljung (1976b), Anderson and Taylor (1979), Lai and Wei (1982), and Rootzén and Sternby (1984).

8.9 PROBLEMS

8G.1. *Local minima:* Consider the matrix $\Gamma_2(\theta)$ defined in Problem 4G.4 and recall that the condition

$$\Gamma_2(\theta) > 0 \tag{8.107a}$$

implies local identifiability of the model structure at θ. Show that this condition together with the condition

$$\Phi_{\chi_0}(\omega) > 0, \qquad \forall\omega \tag{8.107b}$$

[$\Phi_{\chi_0}(\omega)$ is the spectrum of $\chi_0(t)$ defined in (8.14a)] implies that

$$\bar{E}\psi(t,\theta)\psi^T(t,\theta) > 0$$

where $\psi(t,\theta)$ as usual is $(d/d\theta)\hat{y}(t|\theta)$. Show also that if $\varepsilon(t,\theta_0) = e_0(t)$ is white noise and if

$$\bar{E}\ell'(e_0(t)) = 0, \qquad \bar{E}\ell''(e_0(t)) > 0$$

then the conditions (8.107) (at $\theta = \theta_0$) imply that $\bar{V}(\theta) = \bar{E}\ell(\varepsilon(t,\theta))$ has a strict local minimum at $\theta = \theta_0$.

8G.2. Suppose that the transfer-function model set $\{G(q,\theta)\}$ consists of nth-order linear transfer functions:

$$G(q,\theta) = \frac{b_1 q^{-1} + \cdots + b_n q^{-n}}{1 + a_1 q^{-1} + \cdots + a_n q^{-n}}$$

And suppose that the input consists of n sinusoids of different frequencies, ω_i, $i = 1, \ldots, n$, $0 < \omega_i < \pi$.

(a) Suppose that the noise model is independently parametrized. Show that the limit model $G^*(e^{i\omega})$ fits exactly to the true system $G_0(e^{i\omega})$ at the frequencies in question. It is consequently the same result as if we applied frequency analysis with these input frequencies.

(b) Suppose that the noise model has parameters in common with $G(q,\theta)$, but that the system is noise-free: $\Phi_v(\omega) \equiv 0$. Show that the result under part (a) still holds.

8G.3. *Consistency of the maximum likelihood method:* Suppose that the conditions of Theorem 8.5, except (8.52) and (8.53), hold. Let

$$\ell(x) = -\log f_e(x)$$

where $f_e(x)$ is the PDF of $e_0(t)$. Show that (8.52) holds and that the theorem holds even if (8.53) is not satisfied.

8G.4. Consider the frequency-domain expression (8.61) for the criterion function. Show that an alternative expression for $\Phi_\varepsilon(\omega,\theta)$, valid also for inputs generated by feedback, is

$$\Phi_\varepsilon(\omega,\theta) = \lambda_0 + \frac{\tilde{T}(e^{i\omega},\theta)\Phi_{\chi_0}(\omega)\tilde{T}^T(e^{-i\omega},\theta)}{|H(e^{i\omega},\theta)|^2}$$

with the notation (8.15), where $\Phi_{\chi_0}(\omega)$ is the spectrum of $\chi_0(t)$ in (8.14).

8E.1. Apply Theorem 8.2 to the LS criterion (7.33) and verify the heuristically derived result (7.39).

8E.2. Consider Problem 7E.4. Here the criterion function $\ell(\varepsilon(t,\theta),\theta)$ is not parametrized independently from the parametrization of the predictor. Suppose that the true system is given by

$$x(t+1) = a_0 x(t) + w_0(t)$$

$$y(t) = x(t) + v_0(t)$$

where $\{w_0(t)\}$ and $\{v_0(t)\}$ are independent, white Gaussian noises with variances $Ew_0^2(t) = 0.1$ and $Ev_0^2(t) = 10$, respectively.

(a) Show that there exists a value θ^* in the parametrization (7.122) such that

$$\varepsilon(t,\theta^*) = e_0(t) = \text{the true innovations}$$

(b) Show that the maximum likelihood estimate $\hat{\theta}_N^{\text{ML}}$ does *not* converge to θ^* as N tends to infinity.

(c) Explain the paradox.

8E.3. Consider the output error model structure

$$\hat{y}(t|\theta) = \frac{b}{1+fq^{-1}}u(t-1), \qquad \theta = \begin{bmatrix} b \\ f \end{bmatrix}$$

Suppose the true system is as in Example 8.1:

$$y(t) + a_0 y(t-1) = b_0 u(t-1) + e_0(t) + c_0 e_0(t-1)$$

and that the input is generated as

$$u(t) = -k_0 y(t) + r(t)$$

where $\{r(t)\}$ is white noise and

$$|a_0 + k_0 b_0| < 1$$

Give an expression that characterizes the limit of $\hat{\theta}_N$. Will it be equal to $[a_0\ b_0]^T$? On what point are the assumptions of Theorem 8.4 violated?

8E.4. Consider the model structure

$$y(t) + ay(t-1) = bu(t-1) + e(t), \qquad \theta = \begin{bmatrix} a \\ b \end{bmatrix}$$

Suppose that the true system and input are given as in Problem 8E.3. Let $\hat{\theta}_N$ be the IV estimate of θ with instruments

$$\zeta(t) = \begin{bmatrix} u(t-1) \\ u(t-2) \end{bmatrix}$$

Does $\hat{\theta}_N$ converge to $\theta_0 = \begin{bmatrix} a_0 \\ b_0 \end{bmatrix}$?

8E.5. Consider the expression for Φ_ε of Problem 8G.4. Suppose that the feedback can be described as

$$u(t) = K(q)e(t) + w(t)$$

Show that the error spectrum can then be written as

$$\lambda_0 + \frac{\{|\tilde{G}(e^{i\omega},\theta)|^2 \Phi_w(\omega) + |\tilde{G}(e^{i\omega},\theta)K(e^{i\omega}) + \tilde{H}(e^{i\omega},\theta)|^2 \lambda_0\}}{|H(e^{i\omega},\theta)|^2}$$

8T.1. Consider a quasi-stationary sequence $\{u(t)\}$ and a uniformly stable family of filters $\{G(q,\theta), \theta \in D_{\mathcal{M}}\}$. Let

$$z(t,\theta) = G(q,\theta)u(t)$$

Strengthen Theorem 2.2 to hold as follows: For each $\theta \in D_{\mathcal{M}}$, $\{z(t,\theta)\}$ is a quasi-stationary sequence and

$$\sup_{\theta \in D_{\mathcal{M}}} \left| \frac{1}{N} \sum_{t=1}^{N} z^2(t,\theta) - R_\theta(0) \right| \to 0, \qquad \text{as } N \to \infty$$

Here

$$R_\theta(0) = \bar{E} z^2(t,\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |G(e^{i\omega},\theta)|^2 \Phi_u(\omega) d\omega$$

Use this result to give a probabilistic-free counterpart of Theorem 8.2: For any quasi-stationary deterministic sequences $\{y(t)\}$ and $\{u(t)\}$,

$$\hat{\theta}_N \to \arg \min_{\theta \in D_{\mathcal{M}}} \bar{E}\varepsilon^2(t,\theta)$$

where

$$\bar{E}\varepsilon^2(t,\theta) = \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} \varepsilon^2(t,\theta)$$

See Ljung (1985e) for a related discussion.

8T.2. Consider the situation of (8.102). Let the true system be given by (8.93) and suppose that the filter choices are as follows:

$$L(q) = K(q) = 1$$

$$N(q) = A_0(q)$$

$$M(q) = B_0(q)$$

And $\Phi_u(\omega) > 0$ for all ω. Show that $R(\rho)$ is positive definite (and hence nonsingular) for this choice of ρ.

8T.3. Let the system be given by

$$\mathcal{S}: \quad y(t) = G_0(q)u(t) + H_0(q)e_0(t)$$

and an underlying model by

$$\mathcal{M}: \quad y(t) = G(q,\theta)u(t) + H(q,\theta)e(t) \tag{8.108}$$

Suppose that

$$D_T(\mathcal{S},\mathcal{M}) = \{\theta_0\}$$

and that the data are informative. Now let

$$\hat{y}(t|t-k;\theta)$$

be the k-step-ahead predictor computed from (8.108), and let $\hat{\theta}_N$ be determined by

$$\hat{\theta}_N = \arg\min \frac{1}{N}\sum_{t=1}^{N} [y(t) - \hat{y}(t|t-k;\theta)]^2$$

Show that $\hat{\theta}_N \to \theta_0$ as $N \to \infty$, provided there is no output feedback in the generation of the input. What happens if there is feedback? What happens if there is feedback, but there is a k-step time delay between input and output? (*Hint:* Note that the k-step-ahead predictor is a special case of the general linear model so that Theorem 8.2 is applicable. Try to copy the technique of Theorem 8.3.)

8D.1. Show that if

$$\sup_{-1 \le x \le 1} |f_N(x) - \bar{f}(x)| \to 0$$

and

$$x_N = \arg\min_{-1 \le x \le 1} f_N(x)$$

then

$$x_N \to \arg\min \bar{f}(x)$$

8D.2. Show that (8.20) follows from (8.18), (8.2), and (8.5).

8D.3. To show that it is not sufficient to require $\ell(x)$ to be increasing with $|x|$ in Theorem 8.5, consider the following counterexample:

$$\ell(x) = \begin{cases} 0, & x = 0 \\ 2|x|, & 0 \le |x| \le \frac{1}{2} \\ 1, & |x| \ge \frac{1}{2} \end{cases}$$

$$\varepsilon(t,\theta_0) = e_0(t) = \begin{cases} +1, & \text{with probability } 1/2 \\ -1, & \text{with probability } 1/2 \end{cases}$$

Suppose there exists a value $\tilde{\theta}$ such that $\hat{y}(t|\tilde{\theta}) - \hat{y}(t|\theta_0)$ has the following distribution:

$$\begin{cases} +1, & \text{with probability } 1/2 \\ -1, & \text{with probability } 1/2 \end{cases}$$

Check that

(a) Condition (8.52) of Theorem 8.5 is satisfied, but not (8.53).

(b) $E\varepsilon^2(t, \tilde{\theta}) > E\varepsilon^2(t, \theta_0)$.

(c) $E\ell(\varepsilon(t, \tilde{\theta})) < E\ell(\varepsilon(t, \theta_0))$ so that $\hat{\theta}_N \not\to \theta_0$.

8C.1. Consider Problem 7C.2. Draw a mesh plot of $\bar{V}(\theta)$ in the two cases discussed there.