BIRKHÄUSER

Systems & Control: Foundations & Applications

Le Yi Wang
G. George Yin
Ji-Feng Zhang
Yanlong Zhao

System Identification with Quantized Observations

Birkhäuser
Boston • Basel • Berlin

Le Yi Wang
Department of Electrical
and Computer Engineering
Wayne State University
Detroit, MI 48202
USA
lywang@ece.eng.wayne.edu

G. George Yin
Department of Mathematics
Wayne State University
Detroit, MI 48202
USA
gyin@math.wayne.edu

Ji-Feng Zhang
Key Laboratory of Systems and Control
Academy of Mathematics
and Systems Science
Chinese Academy of Sciences
Beijing 100190
China
jif@iss.ac.cn

Yanlong Zhao
Key Laboratory of Systems and Control
Academy of Mathematics
and Systems Science
Chinese Academy of Sciences
Beijing 100190
China
ylzhao@amss.ac.cn

ISBN 978-0-8176-4955-5 e-ISBN 978-0-8176-4956-2
DOI 10.1007/978-0-8176-4956-2

Library of Congress Control Number: 2010927909

Mathematics Subject Classification (2010): Primary: 93B30, 93E12, 93C55; Secondary: 60B10, 62L20

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To my wife, Hong, for her ability to find joy in everything and share it with me, from French cooking to pilates, propofol controlling, or just doing nothing; and hopefully by a long stretch of that ability, in reading this book. And to my daughters, Jackie and Abby, for their support, encouragement, and sympathy after failing to find interesting topics in this book

Le Yi Wang

For Uncle Peter & Aunt Nancy and Uncle Richard & Aunt Helen, who helped me come to the U.S. and launch my academic career, with eternal gratitude

George Yin

To my wife, Ai Ling, and my son, Chao, for giving me a wonderful family full of harmony and happiness, like a peaceful harbor in a turbulent sea that makes me safe and sound

Ji-Feng Zhang

To my wife, Ying, for the magic of turning everything into happiness and for her support when I considered the topics included in this book during my overseas visits. To my family in my hometown for their selfless love, and, especially, in memory of my grandfather

Yanlong Zhao

Contents

Preface

This book concerns the identification of systems in which only quantized output observations are available, due to sensor limitations, signal quantization, or coding for communications. Although there are many excellent treaties in system identification and its related subject areas, a systematic study of identification with quantized data is still in its early stage. This book presents new methodologies that utilize quantized information in system identification and explores their potential in extending control capabilities for systems with limited sensor information or networked systems.

The book is an outgrowth of our recent research on quantized identification; it offers several salient features. From the viewpoint of targeted plants, it treats both linear and nonlinear systems, and both time-invariant and time-varying systems. In terms of noise types, it includes independent and dependent noises, stochastic disturbances and deterministic bounded noises, and noises with unknown distribution functions. The key methodologies of the book combine empirical measures and information-theoretic approaches to cover convergence, convergence rate, estimator efficiency, input design, threshold selection, and complexity analysis. We hope that it can shed new insights and perspectives for system identification.

The book is written for systems theorists, control engineers, applied mathematicians, as well as practitioners who apply identification algorithms in their work. The results presented in the book are also relevant and useful to researchers working in systems theory, communication and computer networks, signal processing, sensor networks, mobile agents, data fusion, remote sensing, tele-medicine, etc., in which noise-corrupted quantized data need to be processed. Selected materials from the book may also be used in a graduate-level course on system identification.

This book project could not have been completed without the help and encouragement of many people. We first recognize our institutions and colleagues for providing us with a stimulating and supportive academic environment. We thank the series editor Tamer Başar for his encouragement and consideration. Our thanks also go to Birkhäuser editor Tom Grasso for his assistance and help, and to the production manager, and the Birkhäuser professionals for their work in finalizing the book. Our appreciation also goes to three anonymous reviewers, who read an early version of the book and offered many insightful comments and suggestions. During the years of study, our research has been supported in part by the National Science Foundation and the National Security Agency of the United States, and by the National Natural Science Foundation of China. Their support is greatly appreciated. We are deeply indebted to many researchers in the field for insightful discussions and constructive criticisms, and for enriching us with their expertise and enthusiasm. Most importantly, we credit our families for their unconditional support and encouragement even when they question our wisdom in working so tirelessly on mathematics symbols.

Detroit Le Yi Wang
Detroit George Yin
Beijing Ji-Feng Zhang
Beijing Yanlong Zhao

Conventions

This book uses a chapter-indexed numbering system for equations, theorems, etc., divided into three groups: (1) equations; (2) definitions, theorems, lemmas, corollaries, examples, propositions, and remarks; (3) assumptions. Each group uses its own number sequencing. For example, equation (3.10) indicates the tenth equation in Chapter 3. Similarly, group 2 entries are sequenced as Definition 4.1, Theorem 4.2, Corollary 4.3, Remark 4.4, Example 4.5 in Chapter 4. Assumptions are marked with the prefix "A" such as (A6.1), which indicates the first-listed assumption in Chapter 6.

In this book, the subscript is mainly used as a time index or iteration number for a sequence, such as y_k for signals at time k, a_l for the lth value of the system impulse response, and θ_N for the estimate at the Nth iteration. We limit the usage of superscripts whenever possible to avoid confusion with polynomial powers, or double subscripts to reduce convoluted notation, and will confine them in local sections when they must be used. The further dependence of a symbol on other variables such as vector or matrix indices, parameters, data length, etc., will be included in parentheses. For example, for a vector process y_k, $y_k^{\{i\}}$ denotes its ith element; $M(i,j)$ or $M^{\{i,j\}}$ represents the element at the ith row and jth column of a matrix M; $e_N(\theta,\mu)$ or $M(i,j;\theta,\mu)$ indicates their dependence on parameters θ and μ, although such a dependence will be suppressed when it becomes clear from the context. For a quick reference, in what follows we provide a glossary of symbols used throughout the book.

Glossary of Symbols

A'	transpose of a matrix or a vector A
A^{-1}	inverse of a matrix A
$\mathrm{Ball}_p(c, r)$	closed ball of center c and radius r in the l^p norm
$\mathbb{C}$	space of complex numbers
C_i	ith threshold of a quantized sensor
CR lower bound	Cramér–Rao lower bound
$E\xi$	expectation of a random variable ξ
$F(\cdot)$	probability distribution function
$\mathcal{F}$	σ-algebra
$\{\mathcal{F}_t\}$	filtration $\{\mathcal{F}_t, t \geq 0\}$
$G(v)$	componentwise operation of a scalar function G on a vector $v = [v^{\{1\}}, \ldots, v^{\{n\}}]'$, $G(v) = \left[G(v^{\{1\}}) \right.$, $\ldots$, $\left. G(v^{\{n\}})\right]'$
$G^{-1}(v)$	componentwise inverse of a scalar invertible function G on a vector v: $G^{-1}(v) = \left[G^{-1}(v^{\{1\}}) \right.$, $\ldots$, $\left. G^{-1}(v^{\{n\}})\right]'$
$H(e^{i\omega})$	scalar or vector complex function of ω
I	identity matrix of suitable dimension
I_A	indicator function of the set A
$\mathbb{N}$	set of natural numbers
ODE	ordinary differential equation
QCCE	quasi-convex combination estimate
$O(y)$	function of y satisfying $\sup_{y \neq 0} \lvert O(y)\rvert / \lvert y\rvert < \infty$
$\mathbb{R}^n$	n-dimensional real-valued Euclidean space

$\mathrm{Rad}_p(\Omega)$	radius of an uncertainty set Ω in l^p
$\mathcal{S}$	binary-valued or quantized sensor
T	Toeplitz matrix
$\mathbb{Z}_+$	set of positive integers
a^+	$= \max\{a, 0\}$ for a real number a
a^-	$= -\max\{-a, 0\}$ for a real number a
a.e.	almost everywhere
a.s.	almost sure
$\mathrm{diag}(A^1, \ldots, A^l)$	diagonal matrix of blocks $A^1, \ldots, A^l$
$f(\cdot)$	probability density function $f(x) = dF(x)/dx$
g_x or $\nabla_x g$	gradient of a function g with respect to (w.r.t.) x
$\mathtt{i}$	pure imaginary number with $\mathtt{i}^2 = -1$
i.i.d.	independent and identically distributed
$\ln x$	natural logarithm of x
$\log x$ or $\log_2 x$	base 2 logarithm of x
$o(y)$	function of y satisfying $\lim_{y\to 0} o(y)/\lvert y\rvert = 0$
q	one-step delay operator: $qx_k = x_{k-1}$
$s_k = \mathcal{S}(y_k)$	output of a sensor, either scalar or vector
$\mathrm{tr}(A)$	trace of the matrix A
$v^{\{i\}}$	ith component of the vector v
w.p.1	with probability one
$\lceil x \rceil$	ceiling function: the smallest integer that is $\geq x$
$\lfloor x \rfloor$	floor function: the largest integer that is $\leq x$
$\lVert x\rVert_p$	l^p norm of a sequence of real numbers $x = \{x_k; k \in \mathbb{N}\}$
Φ_N	$= [\phi_0, \ldots, \phi_{N-1}]'$, regression matrix at iteration N
$(\Omega, \mathcal{F}, P)$	probability space
$\varrho(c, r)$	neighborhood about c of radius r
θ	system parameter (column) vector of the modeled part
$\tilde{\theta}$	system parameter vector of the unmodeled dynamics, usually infinite dimensional
θ_N	estimate of θ at iteration step N
ϕ_k	regression (column) vector of θ at time k
$\tilde{\phi}_k$	regression vector of $\tilde{\theta}$ at time k, usually infinite dimensional
$:=$ or $\stackrel{\text{def}}{=}$	defined to be
$\mathbb{1}$	column vector with all elements equal to one
$\square$	end of a proof
$\lvert \cdot \rvert$	absolute value of a scalar or the Euclidean norm
$\lVert \cdot \rVert$	of a vector largest singular value of a matrix

Part I

Overview

1

Introduction

This book studies the identification of systems in which only quantized output observations are available. The corresponding problem is termed *quantized identification*.

Sampling and quantization were initially introduced into control systems as part of the computer revolution when controllers became increasingly implemented in computers. When computers had very limited memory and low speeds, errors due to sampling and quantization were substantial. Dramatic improvements in computer memory, speed, precision, and computational power made these considerations more an academic delicacy than a practical constraint. This seems to be the case even for wired and dedicated computer networks for which fiber optic networks can carry large quantities of data with lightning speed.

The recent explosive development in computer and communication networks has ushered in a new era of information processing. Thousands, even millions, of computers are interconnected using a heterogeneous network of wireless and wired systems. Due to fundamental limitations on bandwidth resources in wireless communications and large numbers of customers who share network resources, bandwidths have become a bottleneck for nearly all modern networks. Similar concerns arise in special-purpose networks such as smart sensors, MEMS (micro electromechanical systems), sensor networks, mobile agents, distributed systems, and remote-controlled systems, which have very limited power for communications and whose data-flow rates carry significant costs and limitations. These developments have made the issue of sampling and quantization once again fundamental for theoretical development and practical applications [13, 61, 80, 81, 82].

L.Y. Wang et al., *System Identification with Quantized Observations*, Systems & Control: Foundations & Applications, DOI 10.1007/978-0-8176-4956-2_1,

Consider, for example, computer information processing of a continuous-time system whose output is sampled with a sampling rate of N Hz and quantized with a precision word-length of B bits. Its output observations carry the data-flow rate of NB bits per second (bps). For a typical 16-bit precision and 2-KHz sampling rate, a $32K$-bps bandwidth of data transmission resource is required, on observations of one output alone. This is a significant resource demand, especially when wireless communications of data are involved.

Additionally, quantized sensors are commonly employed in practical systems [10, 42, 49, 73, 90, 98, 99]. Usually they are more cost-effective than regular sensors. In many applications, they are the only ones available during real-time operations. There are numerous examples of binary-valued or quantized observations such as switching sensors for exhaust gas oxygen, ABS (anti-lock braking systems), and shift-by-wire in automotive applications; photoelectric sensors for positions, and Hall-effect sensors for speed and acceleration for motors; chemical process sensors for vacuum, pressure, and power levels; traffic condition indicators in the asynchronous transmission mode (ATM) networks; and gas content sensors (CO, CO_2, H_2, etc.) in the gas and oil industries. In medical applications, estimation and prediction of causal effects with dichotomous outcomes are closely related to binary sensor systems. The following examples represent some typical scenarios.

1.1 Motivating Examples

ATM ABR Traffic Control

An ATM network, depicted in Figure 1.1, consists of sources, switches, and destinations. Due to variations in other higher-priority network traffic such as constant bit rate (CBR) and variable bit rate (VBR), an available bit rate (ABR) connection experiences significant uncertainty on the available bandwidth during its operation. A physical or logical buffer is used in a switch to accommodate bandwidth fluctuations. The actual amount of bandwidth an ABR connection receives is provided to the source using rate-based closed-loop feedback control. One typical technique for providing traffic information is relative rate marking, which uses two fields in the Resource Management (RM) cell—the No Increase (NI) bit and the Congestion Indication (CI) bit. The NI bit is set when the queue reaches length C_1, and the CI bit is set when the queue length reaches C_2 ($C_2 > C_1$).

In this system, the queue length is not directly available for traffic control. The NI and CI bits indicate merely that it takes values in one of the three uncertainty sets $[0, C_1]$, $(C_1, C_2]$, and (C_2, ∞). This can be represented by two binary sensors. It is noted that the desired queue length is usually a value different than C_1 or C_2.

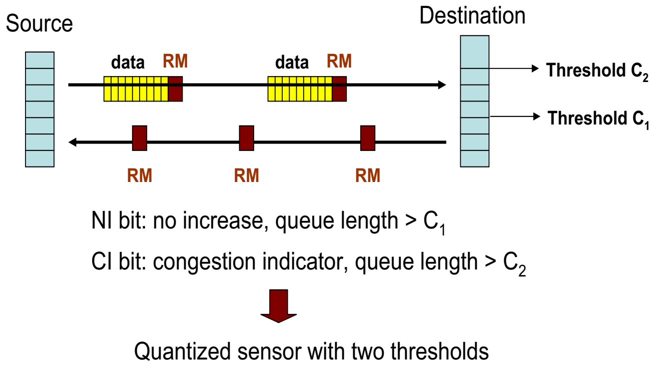

FIGURE 1.1. ATM network control

LNT and Air-to-Fuel Ratio Control with an EGO Sensor

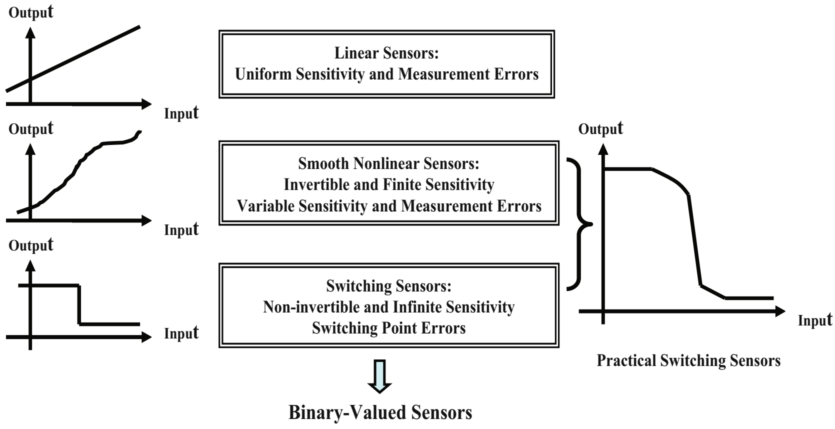

FIGURE 1.2. Sensor types

In automotive and chemical process applications, oxygen sensors are widely used for evaluating gas oxygen contents. Inexpensive oxygen sensors are switching types that change their voltage outputs sharply when excess oxygen in the gas is detected; see Figure 1.2. In particular, in automotive emission control, the exhaust gas oxygen sensor (EGO or HEGO) will switch its outputs when the air-to-fuel ratio in the exhaust gas crosses the stoichiometric value. To maintain the conversion efficiency of the three-way catalyst or to optimize the performance of a lean NOx trap (LNT), it is essential to estimate the internally stored NOx and oxygen. In this case, the switching point of the sensor has no direct bearing on the control target. The idea of using the switching sensor for identification purposes, rather than for control only, can be found in [98, 99, 112].

Identification of Binary Perceptrons

There is an interesting intersection between quantized identification and statistical learning theory in neural networks. Consider an unknown binary perceptron depicted in Figure 1.3 that is used to represent a dynamic relationship:

$$y(t) = \mathcal{S}(w_1x_1 + w_2x_2 + \cdots + w_nx_n - C + d),$$

where C is the known neuron firing threshold, $w_1, \ldots, w_n$ are the weighting coefficients to be learned, and $\mathcal{S}(\cdot)$ is a binary-valued function switching at 0. This learning problem can be formulated as a special case of binary sensor identification without unmodeled dynamics. Traditional neural models, such as the McCulloch–Pitts and Nagumo–Sato models, contain a neural firing threshold that naturally introduces a binary function [13, 38, 42, 73]. Fundamental stochastic neural learning theory studies the stochastic updating algorithms for neural parameters [94, 95, 96].

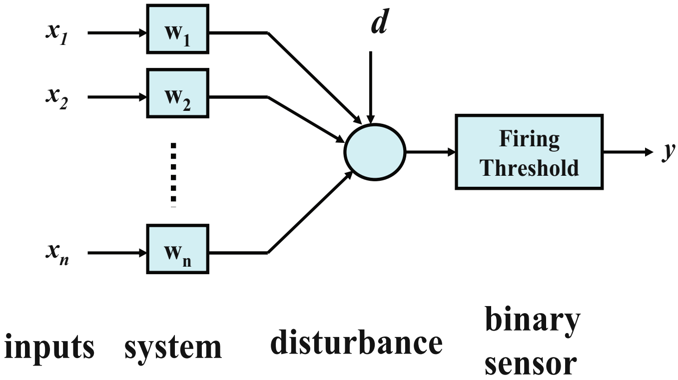

FIGURE 1.3. Artificial neural networks

Networked Systems

In a networked system, see Figure 1.4, signals must be transmitted through communication channels. Since communications make quantization mandatory, it is essential to understand identification with quantized observations. Unlike physical sensors whose characteristics such as switching thresholds cannot be altered during identification experiments, quantization for communication may be viewed generally as a partition of the signal range into a collection of subsets. Consequently, quantization thresholds may be selected to reduce identification errors, leading to the problems on threshold selection. Furthermore, source coding and channel coding after quantization play an important role in identification error characterization when communication channels are noisy.

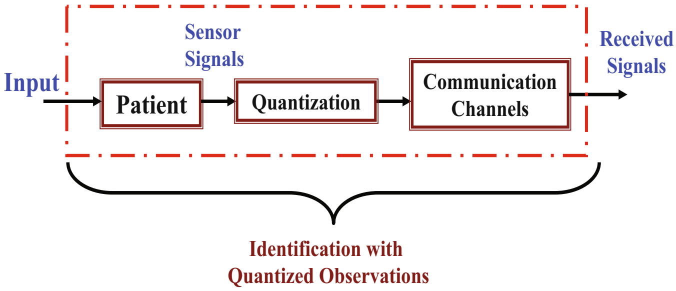

FIGURE 1.4. Communication networks

1.2 System Identification with Quantized Observations

The most common example of quantization is to divide the output range into equal-length intervals. This book considers a general class of quantized observations that allows a partition of the output range into a finite collection of subsets. The subsets may have unequal lengths or be unbounded, may be fixed due to sensor configuration limitations, or may be design variables such as quantization or coding in communication systems. This subject is of importance in understanding the modeling capability for systems with limited sensor information, establishing relationships between communication resource limitations and identification complexity, and studying sensor networks [1, 10, 13, 34, 61, 80, 81, 90, 112, 115].

The use of quantized observations introduces substantial difficulties since only very limited information is available for system modeling, identification, and control. Since switching sensors are nonlinear components, studies of their roles and impact on systems are often carried out in nonlinear system frameworks, such as sliding mode control, describing function analysis, switching control, hybrid control, etc. In these control schemes, the switching thresholds of the sensors are directly used to define a control target. However, their fundamental impact on system modeling and identification is a relatively new area. This book presents recent developments on the inherent consequences of using quantized observations in system identification as well as methods and algorithms that use quantized observations effectively to extend control capabilities.

The main motivation is embodied in many applications in which modeling of such systems is of great importance in performing model predictive control, model-based diagnosis, outcome predictions, optimal control strategy development, control adaptation, etc. When inputs can be arbitrarily selected within certain bounds and outputs are measured by regular sensors, system identification problems have been studied extensively in the traditional setup under the frameworks of either stochastic systems or worst-case identification. The issues of identification accuracy, convergence, model complexity, time complexity, input design, persistent excitation, identification algorithms, etc. have been pursued by many researchers. A vast literature is now available on this topic; see [17, 55, 62, 67], among others.

It should be clarified that the treatments of this book will be meaningful only for the application problems in which quantization levels carry a substantial system cost or operational penalty. If an application can use cheaper sensors of high precision and data-flow rates do not carry a cost, traditional system identification methods will suffice. On the other hand, when a sensor of lower precision is cheaper than a higher-precision sensor, it is important to understand what the performance penalty will be if the cheaper sensor is used. Similarly, when communication bandwidths are limited, the reduction of quantization levels will save communication resources. Intrinsically, one may choose to use coarse quantization (lower space complexity) so that more data points can be transmitted (higher time complexity) with the same bandwidth resource demand. This tradeoff between sampling rates and quantization accuracy is a fundamental issue in complexity analysis for understanding the impact of communication channels on system performance.

Some fundamental issues emerge when the output observation is quantized: How accurately can one identify the parameters of the system? How fast can one reduce uncertainty on model parameters? What are the optimal inputs for fast identification? What are the conditions that ensure the convergence of the identification algorithms? What are the impacts of unmodeled dynamics and disturbances on identification accuracy and time complexity? In contrast to classical system identification, answers to these familiar questions under switching sensors differ substantially from the traditional setup.

This book demonstrates that quantized observations increase time complexity significantly; the optimal inputs differ from those in traditional identification; identification characteristics depart significantly between stochastic and deterministic noise representations; and unmodeled dynamics have a fundamental influence on identification accuracy of the modeled part. In contrast to traditional system identification, in which the individual merits of stochastic versus worst-case frameworks are sometimes debated, it is beneficial to combine these two frameworks in quantized identification problems.

1.3 Outline of the Book

This book is organized into five parts as follows: I. Overview; II. Stochastic Methods for Linear Systems; III. Deterministic Methods for Linear Systems; IV. Identification of Nonlinear and Switching Systems; V. Complexity Analysis.

Part I is an overview that provides motivational applications for system identification with quantized observations in Chapter 1 and that introduces the common systems settings for the entire book in Chapter 2. After a general introduction of the problems in Chapter 1, generic system settings are formulated in Chapter 2. Special cases of systems are then delineated, such as gain systems, finite impulse-response systems, rational systems, and nonlinear systems. The main issues of system identification are further explained, including typical inputs and signal ranks. System uncertainties considered in this book consist of unmodeled dynamics for linear dynamics, model mismatch for nonlinear functions, and disturbances in either a stochastic or deterministic worst-case description. Identification in different system configurations is outlined, in which distinct issues arising from open-loop and closed-loop systems, and input and actuator noises are further discussed.

In this book, unmodeled dynamics are always described as a bounded but unknown uncertainty. In contrast, disturbances are modeled either in a stochastic framework, or as an unknown but bounded uncertainty. Since these two frameworks require vastly different input design, employ distinct analysis methods, and entail diversified convergence properties, they are presented in their elementary but fundamental forms in Parts II and III, respectively.

Part II covers stochastic methods for linear systems with quantized observations. It presents identification algorithms, properties of estimators, and various convergence results in a stochastic system framework. Chapter 3 introduces the main methodology of empirical measures and derives convergence properties, including strong convergence, convergence in distribution, and mean-square convergence. When noise is modeled as a stochastic process and the system is subject to unmodeled dynamics, it becomes mandatory to deal with a combined framework, which is investigated in Chapter 4. Upper and lower error bounds are derived.

When dealing with a complicated system structure, a fundamental idea is to reduce the identification of its parameters to a finite set of identification of gains. This is achieved by employing some intrinsic properties of periodic inputs and invertible mappings. In Chapter 3, we show how a full-rank periodic input can reduce the problem of identifying a finite impulse-response system to a number of core identification problems of gains. When the system is rational, the problem becomes much more difficult. We show in Chapter 5 that this difficulty can be overcome by full-rank periodic inputs when the rational model is co-prime.

The convergence and efficiency of estimators in quantized identification are studied in Chapter 6. When the observation becomes quantized with a finite number of thresholds, an algorithm, termed optimal quasi-convex combination estimation (optimal QCCE), is introduced to derive an estimate from multiple thresholds. The resulting estimate is shown to be asymptotically efficient by comparing its convergence speed to the Cramér–Rao (CR) lower bound.

The utility of full-rank periodic inputs is further investigated in Chapter 7 in typical system configurations such as cascade and feedback connections. It is revealed that periodicity and full-rankness of a signal are two fundamental input properties that are preserved after the signal passes through a stable system with some mild constraints. Consequently, it becomes clear that the input design, identification algorithms, and convergence properties are inherently invariant under open-loop and closed-loop configurations. Furthermore, we present in Chapter 8 the joint identification of system parameters, unknown thresholds, and unknown noise distribution functions. The main idea is to use signal scaling to excite further information on sensor thresholds and noise distribution functions.

Part III is concerned with deterministic methods for linear systems. Here the emphasis is shifted to the deterministic framework for disturbances. Under this framework, noise is modeled as unknown but bounded. Our exploration starts in Chapter 9 with the case of binary-valued observations. Input design that utilizes the idea of bisection is shown to reduce uncertainty exponentially. This idea is employed when both observation noise and unmodeled dynamics are present. Explicit error bounds are derived. Chapter 10 considers the case of quantized observations. The utility of multiple thresholds in accelerating convergence speed is investigated.

Part IV concentrates on the identification of nonlinear and switching systems. The first concentration is on Wiener and Hammerstein systems in which the nonlinearity is confined to be memoryless. The algorithms for identifying such nonlinear systems closely follow the ideas of signal scaling in Chapter 8 to separate the identification of the linear dynamics and nonlinear function and to extract information on the nonlinear part. This is especially apparent in Chapter 11 for Wiener systems. Hammerstein systems are treated in Chapter 12. Although there are similarities between Wiener and Hammerstein systems, input design is more stringent in Hammerstein systems. Some new concepts of input ranks are introduced. Systems with switching parameters are discussed in Chapter 13. In such systems, parameters are themselves Markov chains. Two essential cases are included. In the first case, parameters switch their values much faster than identification convergence speeds. Consequently, it is only feasible to identify the average of the switching parameters. On the other hand, if the parameter jumps occur infrequently with respect to identification speeds, parameter tracking by identification algorithms can be accomplished. Algorithms, convergence, and convergence rates toward an irreducible uncertainty set are established.

Part V explores fundamental complexity issues in system identification with quantized observations. The main tool is the asymptotic efficiency that defines the impact of observation time complexity (data length) and space complexity (number of thresholds) on identification accuracy. The tradeoff between time complexity and space complexity points to a broad utility in threshold selection, optimal resource allocations, and communication quantization design. These discussions are contained in Chapter 14. This understanding is further utilized to study the impact of communication channels on system identification in Chapter 15. The concept of the Fisher information ratio is introduced.

In addition to the aforementioned chapters and an extensive list of references at the end of the book, each chapter (except for Chapter 1) has a section of notes in which historical remarks, developments of related work and references, and possible future study topics are presented and discussed.

2

System Settings

This chapter presents basic system structures, sensor representations, input types and characterizations, system configurations, and uncertainty types for the entire book. This chapter provides a problem formulation, shows connections among different system settings, and demonstrates an overall picture of the diverse system identification problems that will be covered in this book. Other than a few common features, technical details are deferred to later chapters.

Section 2.1 presents the basic system structure and its special cases of FIR (finite impulse response), IIR (infinite impulse response), rational, and nonlinear systems that will be discussed in detail in later chapters. Quantized observations and their mathematical representations are described in Section 2.2. Essential properties of periodic input signals that are critical for quantized identification are established in Section 2.3. When a system is embedded in a larger configuration, its input and output are further confined by the system structure, introducing different identification problems. Section 2.4 shows several typical system configurations and their corresponding unique features in system identification. There are many types of uncertainties that can potentially impact system identification. These are summarized in Section 2.5.

L.Y. Wang et al., *System Identification with Quantized Observations*, Systems & Control: Foundations & Applications, DOI 10.1007/978-0-8176-4956-2_2,

2.1 Basic Systems

The basic system structure under consideration is a single-input–single-output stable system in its generic form

$$y_k = G(U_k, \theta) + \Delta(U_k, \widetilde{\theta}) + d_k, \quad k = 0, 1, 2, \ldots, \tag{2.1}$$

where $U_k = \{u_j, 0 \le j \le k\}$ is the input sequence up to the current time k, $\{d_k\}$ is a sequence of random variables representing disturbance, θ is the vector-valued parameter to be identified, and $\widetilde{\theta}$ represents the unmodeled dynamics. All systems will assume zero initial conditions, which will not be mentioned further in this book. We first list several typical cases of the system in (2.1).

1. **Gain Systems:**
$$y_k = a u_k + d_k.$$
Hence, $\theta = a$, $G(U_k, \theta) = a u_k$, and $\Delta(U_k, \widetilde{\theta}) = 0$.

2. **Finite Impulse Response (FIR) Models:**
$$y_k = a_0 u_k + \cdots + a_{n_0-1} u_{k-n_0+1} + d_k.$$
This is usually written in a regression form
$$G(U_k, \theta) = a_0 u_k + \cdots + a_{n_0-1} u_{k-n_0+1} = \phi_k' \theta,$$
where $\theta = [a_0, \ldots, a_{n_0-1}]'$ is the unknown parameter vector and $\phi_k' = [u_k, \ldots, u_{k-n_0+1}]$ is the regressor. In this case, the model order is n_0, which is sometimes used as a measure of model complexity. Again, $\Delta(U_k, \widetilde{\theta}) = 0$.

3. **Infinite Impulse Response (IIR) Models:**
$$y_k = \sum_{n=0}^{\infty} a_n u_{k-n} + d_k,$$
where the system parameters satisfy the bounded-input–bounded-output (BIBO) stability constraint
$$\sum_{n=0}^{\infty} |a_n| < \infty.$$
For system identification, this model is usually decomposed into two parts:
$$\sum_{n=0}^{n_0-1} a_n u_{k-n} + \sum_{n=n_0}^{\infty} a_n u_{k-n} = \phi_k' \theta + \widetilde{\phi}_k' \widetilde{\theta}, \tag{2.2}$$

where $\theta = [a_0, \ldots, a_{n_0-1}]'$ is the modeled part and $\widetilde{\theta} = [a_{n_0}, a_{n_0+1}, \ldots]'$ is the unmodeled dynamics, with corresponding regressors

$$\phi_k' = [u_k, \ldots, u_{k-n_0+1}] \quad \text{and}$$
$$\widetilde{\phi}_k' = [u_{k-n_0}, u_{k-n_0-1}, \ldots],$$

respectively. In this case, the model order is n_0. For the selected n_0, we have

$$G(U_k, \theta) = \phi_k' \theta; \quad \Delta(U_k, \widetilde{\theta}) = \widetilde{\phi}_k' \widetilde{\theta}.$$

4. **Rational Transfer Functions:**

$$y_k = G(q, \theta) u_k + d_k. \tag{2.3}$$

Here q is the one-step shift operator $qu_k = u_{k-1}$ and $G(q)$ is a stable rational function[2.1] of q:

$$G(q) = \frac{B(q)}{1 - A(q)} = \frac{b_1 q + \cdots + b_{n_0} q^{n_0}}{1 - (a_1 q + \cdots + a_{n_0} q^{n_0})}.$$

In this case, the model order is n_0 and the system has $2n_0$ unknown parameters $\theta = [a_1, \ldots, a_{n_0}, b_1, \ldots, b_{n_0}]'$. Note that in this scenario, the system output is nonlinear in parameters. To relate it to sensor measurement errors in practical system configurations, we adopt the output disturbance setting in (2.3), rather than the equation disturbance structure in

$$y_k + a_1 y_{k-1} + \cdots + a_{n_0} y_{k-n_0} = b_1 u_{k-1} + \cdots + b_{n_0} u_{k-n_0} + d_k,$$

which is an autoregressive moving average (ARMA) model structure. The ARMA model structure is more convenient for algorithm development. But output measurement noises in real applications occur in the form of (2.3).

5. **Wiener Models:**

$$G(U_k, \theta) = H(G_0(q, \theta_1) u_k, \beta),$$

or in a more detailed expression

$$x_k = \sum_{n=0}^{n_0-1} a_n u_{k-n}, \quad y_k = H(x_k, \beta) + d_k.$$

[2.1] When $G(q, \theta)$ is used in a closed-loop system, it will be allowed to be unstable, but is assumed to be stabilized by the feedback loop.

Here, β is the parameter (column) vector of the output memoryless nonlinear function H and $\theta_1 = [a_0, \ldots, a_{n_0-1}]'$ is the parameter vector of the linear part. The combined unknown parameters are $\theta = [\theta_1', \beta']'$.

6. **Hammerstein Models:**

$$G(U_k, \theta) = G_0(q, \theta_1)H(u_k, \beta)$$

or

$$y_k = \sum_{n=0}^{n_0-1} a_n x_{k-n} + d_k, \quad x_k = H(u_k, \beta).$$

Here, β is the parameter vector of the input memoryless nonlinear function H and $\theta_1 = [a_0, \ldots, a_{n_0-1}]'$ is the parameter vector of the linear part. The combined unknown parameters are $\theta = [\theta_1', \beta']'$.

2.2 Quantized Output Observations

Let us begin with Figure 2.1. The output y_k in (2.1) is measured by a

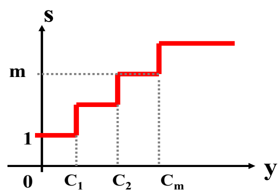

FIGURE 2.1. Quantized observations

sensor of m_0 thresholds $-\infty < C_1 < \ldots < C_{m_0} < \infty$. The sensor can be represented by a set of m_0 indicator functions $s_k = [s_k(1), \ldots, s_k(m_0)]'$, where $s_k(i) = I_{\{-\infty < y_k \leq C_i\}}$, $i = 1, \ldots, m_0$, and

$$I_{\{y_k \in A\}} = \begin{cases} 1, & \text{if } y_k \in A, \\ 0, & \text{otherwise.} \end{cases}$$

In such a setting, the sensor is modeled as m_0 binary-valued sensors with overlapping switching intervals, which imply that if $s_k(i) = 1$, then $s_k(j) = 1$ for $j \geq i$. An alternative representation of the sensor is by defining $\widetilde{s}_k(i) = I_{\{C_{i-1} < y_k \leq C_i\}}$ with $C_0 = -\infty$, and $C_{m_0+1} = \infty$ with the interval (C_{m_0}, ∞). This representation employs distinct switching intervals. Consequently, only one $s_k(i) = 1$ at any k.

Under a quantized sensor of m_0 thresholds, each sample of the signal can be represented by a code of length $\log_2 m_0$ bits. This will be viewed as the space complexity of the signal measurements.

2.3 Inputs

In this book, we use extensively periodic input signals in identification experiments under a stochastic framework. A signal v_k is said to be n_0-periodic if $v_{k+n_0} = v_k$. We first establish some essential properties of periodic signals, which will play an important role in the subsequent development.

Toeplitz Matrices

Recall that an $n_0 \times n_0$ Toeplitz matrix [37] is any matrix with constant values along each (top-left to bottom-right) diagonal. That is, a Toeplitz matrix has the form

$$T = \begin{bmatrix} v_{n_0} & \dots & v_2 & v_1 \\ v_{n_0+1} & \ddots & \ddots & v_2 \\ \vdots & \ddots & \ddots & \vdots \\ v_{2n_0-1} & \dots & v_{n_0+1} & v_{n_0} \end{bmatrix}.$$

It is clear that a Toeplitz matrix is completely determined by its entries in the first row and the first column $\{v_1, \dots, v_{2n_0-1}\}$, which is referred to as the symbol of the Toeplitz matrix.

Circulant Toeplitz Matrices and Periodic Signals

A Toeplitz matrix T is said to be circulant if its symbol satisfies $v_k = v_{k-n_0}$ for $k = n_0 + 1, \dots, 2n_0 - 1$; see [25]. A circulant matrix [57] is completely determined by its entries in the first row $[v_{n_0}, \dots, v_1]$, so we denote it as $T([v_{n_0}, \dots, v_1])$. Moreover, T is said to be a generalized circulant matrix if $v_k = \rho v_{k-n_0}$ for $k = n_0 + 1, \dots, 2n_0 - 1$, where $\rho > 0$, which is denoted by $T(\rho, [v_{n_0}, \dots, v_1])$ and

$$T(\rho, [v_{n_0}, \dots, v_1]) = \begin{bmatrix} v_{n_0} & \dots & v_2 & v_1 \\ \rho v_1 & \ddots & \ddots & v_2 \\ \vdots & \ddots & \ddots & \vdots \\ \rho v_{n_0-1} & \dots & \rho v_1 & v_{n_0} \end{bmatrix}.$$

Definition 2.1. An n_0-periodic signal generated from its one-period values $v = (v_1, \dots, v_{n_0})$ is said to be full rank if $T([v_{n_0}, \dots, v_1])$, the circulant matrix, is full rank.

An important property of circulant matrices is the following frequency-domain criterion.