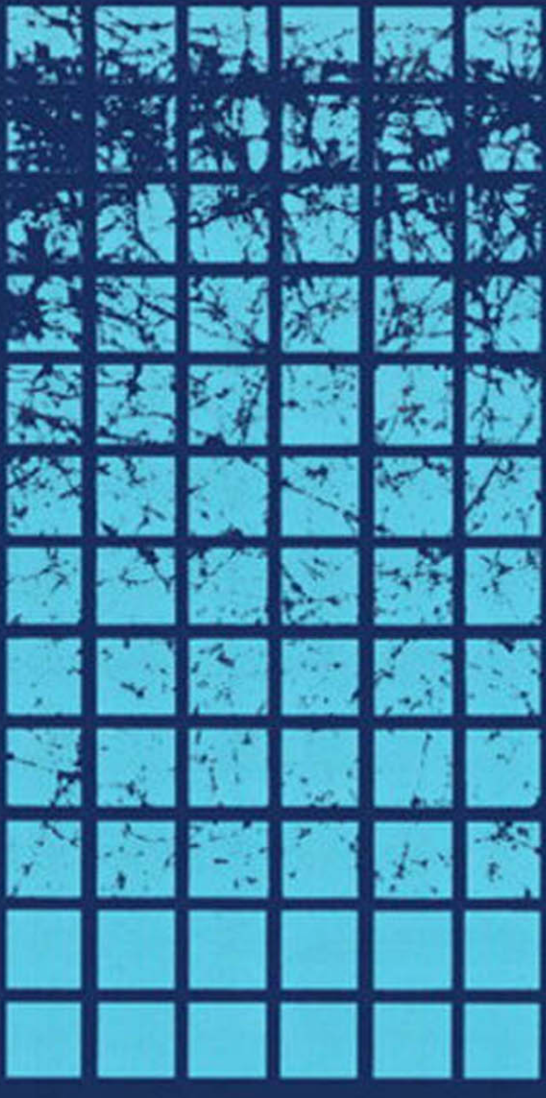
INFORMATION AND SYSTEM SCIENCES SERIES
Thomas Kailath, Series Editor

SYSTEM IDENTIFICATION

second edition

Theory FOR THE User



LENNART
LJUNG

System Identification

Theory for the User

Second Edition

ISBN 0-13-656695-2

90000

9 780136 566953

5E.3 Show (5.56) by induction as follows: Suppose first that there is just one regressor, associated with k attributes. Then there must be k rules in the rule base, for it to be complete, covering all the attributes. Thus (5.56) follows from the assumption (5.54). Now suppose that (5.56) holds for d regressors and that there are K rules. Here $K = k_1 \cdot k_2 \cdots k_d$, with k_j as the number of attributes for regressor j. Now, add another regressor φ_{d+1} with k_{d+1} attributes, subject to (5.54). For the rule base to remain complete, it must now be complemented with $K \cdot k_{d+1}$ new rules, covering the combinations of the previous cases with each attribute of the new regressor. Show the induction step, that (5.56) holds also when φ_{d+1} has been added.

5T.1 Time-continuous bilinear system descriptions are common in many fields (see Mohler, 1973). A model can be written

$$\dot{x}(t) = A(\theta)x(t) + B(\theta)u(t) + G(\theta)x(t)u(t) + w(t) \tag{5.78a}$$

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

$$y(t) = C(\theta)x(t) + e(t), \qquad \text{for } t = kT \tag{5.78b}$$

where $e(t)$ is white Gaussian measurement noise with variance R_2. The input is piecewise constant:

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

Derive an expression for the prediction of $y((k+1)T)$, given u_r and $y(rT)$ for $r \leq k$, based on the model (5.78).

5T.2 Consider the Monod growth model structure

$$\dot{x}_1 = \frac{\theta_1 \cdot x_2}{\theta_2 + x_2} \cdot x_1 - \alpha_1 x_1$$

$$\dot{x}_2 = \frac{-1}{\theta_3} \cdot \frac{\theta_1 \cdot x_2}{\theta_2 + x_2} \cdot x_1 - \alpha_1(x_2 - \alpha_2)$$

$y = [x_1 \quad x_2]^T$ 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 13.

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 (4.50)], 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

$$\overline{E}u(t)u(t-\tau) = R_u(\tau)$$

and

$$\overline{E}u(t)v(t-\tau) \equiv 0 \qquad \text{(open-loop operation)}$$

then according to Theorem 2.2 (expressed in the time domain)

$$\overline{E}y(t)u(t-\tau) = R_{yu}(\tau) = \sum_{k=1}^{\infty} g_0(k)R_u(k-\tau) \tag{6.8}$$

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.

In fact, the most natural way to estimate $g(k)$ when the input is not "exactly white" is to truncate (6.7) at n, and treat it as an n:th order FIR model (4.46) with the parametric (least-squares) methods of Chapter 7. Another way is to filter both inputs and outputs by a prefilter that makes the input as white as possible ("input prewhitening") and then compute the correlation function (6.9) for these filtered sequences.

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 *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 graphical 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

$$\begin{aligned} I_c(N) &= \frac{1}{N}\sum_{t=1}^{N} \alpha|G_0(e^{i\omega})|\cos(\omega t + \varphi)\cos\omega t + \frac{1}{N}\sum_{t=1}^{N} v(t)\cos\omega t \\ &= \alpha|G_0(e^{i\omega})|\frac{1}{2}\frac{1}{N}\sum_{t=1}^{N}[\cos\varphi + \cos(2\omega t + \varphi)] \\ &\quad + \frac{1}{N}\sum_{t=1}^{N} v(t)\cos\omega t \qquad (6.16) \\ &= \frac{\alpha}{2}|G_0(e^{i\omega})|\cos\varphi + \alpha|G_0(e^{i\omega})|\frac{1}{2}\frac{1}{N}\sum_{t=1}^{N}\cos(2\omega t + \varphi) \\ &\quad + \frac{1}{N}\sum_{t=1}^{N} v(t)\cos\omega t \end{aligned}$$

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)gives 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

$$I_c(N) - iI_s(N) = \frac{1}{\sqrt{N}} Y_N(\omega) \tag{6.20}$$