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.
2.1 Basic Systems

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

\[ y_k = G(U_k, \theta) + \Delta(U_k, \tilde{\theta}) + d_k, \quad k = 0, 1, 2, \ldots, \quad (2.1) \]

where \( U_k = \{ u_j, 0 \leq j \leq k \} \) is the input sequence up to the current time \( k \), \( \{ d_k \} \) is a sequence of random variables representing disturbance, \( \theta \) is the vector-valued parameter to be identified, and \( \tilde{\theta} \) 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:**

\[ y_k = au_k + d_k. \]

Hence, \( \theta = a, \ G(U_k, \theta) = au_k, \) and \( \Delta(U_k, \tilde{\theta}) = 0. \)

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, \( \Delta(U_k, \tilde{\theta}) = 0. \)

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:

\[ \sum_{n=0}^{n_0-1} a_n u_{k-n} + \sum_{n=n_0}^{\infty} a_n u_{k-n} = \phi'_k \theta + \tilde{\phi}'_k \tilde{\theta}, \quad (2.2) \]
where \( \theta = [a_0, \ldots, a_{n_0-1}]' \) is the modeled part and \( \tilde{\theta} = [a_{n_0}, a_{n_0+1}, \ldots]' \) is the unmodeled dynamics, with corresponding regressors
\[
\phi_k' = [u_k, \ldots, u_{k-n_0+1}] \quad \text{and} \quad \tilde{\phi}_k' = [u_{k-n_0}, u_{k-n_0-1}, \ldots],
\]
respectively. In this case, the model order is \( n_0 \). For the selected \( n_0 \), we have
\[
G(U_k, \theta) = \phi_k' \theta; \quad \Delta(U_k, \tilde{\theta}) = \tilde{\phi}_k' \tilde{\theta}.
\]

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 \( q \):
\[
G(q) = \frac{B(q)}{1 - A(q)} = \frac{b_1q + \cdots + b_{n_0}q^{n_0}}{1 - (a_1q + \cdots + a_{n_0}q^{n_0})}.
\]

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
\[
y_k + a_1y_{k-1} + \cdots + a_{n_0}y_{k-n_0} = b_1u_{k-1} + \cdots + b_{n_0}u_{k-n_0} + d_k,
\]
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, $\beta$ 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, $\beta$ 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 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 $\tilde{s}_k(i) = I_{\{C_{i-1} < y_k \leq C_i\}}$ with $C_0 = -\infty$, and $C_{m_0+1} = \infty$ with the interval $(C_{m_0}, \infty)$. 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

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

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

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

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

An important property of circulant matrices is the following frequency-domain criterion.
Lemma 2.2. If $T = T(\rho, [v_{n_0}, \ldots, v_1])$ is a generalized circulant matrix, then the eigenvalues of $T$ are $\{\rho \gamma_k, k = 1, \ldots, n_0\}$ and the determinant of $T$ is $\det(T) = \prod_{k=1}^{n_0} \rho \gamma_k$, where $\gamma_k$ is the discrete Fourier transform (DFT) of $v_j \rho^{-(j/n_0)}$, $j = 1, \ldots, n_0$:

$$\gamma_k = \sum_{j=1}^{n_0} v_j \rho^{-\frac{j}{n_0}} e^{-i\omega_k j}, \quad \omega_k = \frac{2\pi k}{n_0}, \quad k = 1, \ldots, n_0.$$

Hence, $T$ is full rank if and only if $\gamma_k \neq 0$, $k = 1, \ldots, n_0$.

Proof. Let

$$M = \begin{bmatrix} 0 & I_{n_0-1} \\ \rho & 0 \end{bmatrix},$$

whose characteristic polynomial is $\lambda^{n_0} - \rho$ and eigenvalues are $\rho^{-(j/n_0)} e^{i\omega_k}$, $k = 1, \ldots, n_0$. Then, $T$ can be represented as $T = \sum_{j=1}^{n_0} v_j M^{n_0-j}$. For $k = 1, \ldots, n_0$, if $x_k$ is the corresponding eigenvector of the eigenvalue $\rho^{-(1/n_0)} e^{i\omega_k}$ for $M$, then

$$Tx_k = \sum_{j=1}^{n_0} v_j M^{n_0-j} x_k = \sum_{j=1}^{n_0} v_j (\rho \rho^{-\frac{j}{n_0}} e^{i\omega_k})^{n_0-j} x_k = \rho \gamma_k x_k.$$

Therefore, $\rho \gamma_k$ is an eigenvalue of $T$ and the expression for $\det(T)$ is confirmed. By hypothesis, $\rho > 0$. Hence, $T$ is full rank if and only if $\gamma_k \neq 0$, $k = 1, \ldots, n_0$.

For the special case when $\rho = 1$, we have the following property.

Corollary 2.3. An $n_0$-periodic signal generated from $v = (v_1, \ldots, v_{n_0})$ is full rank if and only if its discrete Fourier transform $\gamma_k = \sum_{j=1}^{n_0} v_j e^{-i\omega_k j}$ is nonzero at $\omega_k = 2\pi k / n_0$, $k = 1, \ldots, n_0$.

Recall that $\Gamma = \{\gamma_1, \ldots, \gamma_{n_0}\}$ are the frequency samples of the $n_0$-periodic signal $v$. Hence, Definition 2.1 may be equivalently stated as “an $n_0$-periodic signal $v$ is said to be full rank if its frequency samples do not contain 0.” In other words, the signal contains $n_0$ nonzero frequency components.

2.4 System Configurations

The basic system (2.1) is a typical open-loop identification setting in which the input can be selected directly by the user and the output noise is the
only disturbance. Practical systems are far more complicated in which a system to be identified is often a subsystem interconnected in different system configurations. Consequently, the input to the plant may not be directly accessible for design and there may be multiple noise corruptions. In this book, we recognize and treat some of these configurations.

2.4.1 Filtering and Feedback Configurations

Consider the system configurations in Figure 2.2. The filtering configuration is an open-loop system where $M$ is linear, time invariant, and stable but may be unknown. The feedback configuration is a general structure of two-degree-of-freedom controllers where $K$ and $F$ are linear, time invariant, may be unstable, but are stabilizing for the closed-loop system. The mapping from $r$ to $u$ is the stable system $M = K/(1 + PKF)$. When $K = 1$, it is a regulator structure, and when $F = 1$, it is a servo-mechanism or tracking structure. Note that system components $M$, $K$, $F$ are usually designed for achieving other goals and cannot be tuned for identification experiment design.

![Diagram of Filtering and Feedback Configurations](image)

FIGURE 2.2. Typical system configurations

In these configurations, the input $u$ to the plant $P$ may be measured but cannot be directly selected. Only the external input $r$ can be designed.

2.4.2 Systems with Communication Channels

The parameters of the system $G$ in Figure 2.3 are to be identified. Two scenarios of system configuration are considered. System identification with quantized sensors is depicted in Figure 2.3(a) in which the observations on $u_k$ and $s_k$ are used. On the other hand, when sensor outputs of a system are transmitted through a communication channel and observed after transmission, the system parameters must be estimated by observing $u_k$ and $w_k$, as shown in Figure 2.3(b). Since communication channels are subject to channel uncertainties, such as channel noises, the identification of system parameters is influenced by channel descriptions. Furthermore, when the channel contains unknown parameters such as unknown noise distribution functions, they must be incorporated into the overall identification problem.
2.5 Uncertainties

The main purpose of system identification is to reduce uncertainty on the system by using information from input-output observations. There are multiple sources of uncertainty in system configuration, modeling, and environments that will have a significant bearing on system identification. In addition to the unknown model parameters that are to be identified, we list below some of the uncertainties to be considered in this book.

2.5.1 System Uncertainties: Unmodeled Dynamics

In practical applications, a dynamic system is usually infinite dimensional. For cost reduction on system analysis, design, and implementation, it is desirable to use a low-order model to represent the system. A typical case is the IIR system (2.2)

\[
\sum_{n=0}^{n_0-1} a_n u_{k-n} + \sum_{n=n_0}^{\infty} a_n u_{k-n} = \phi_k' \hat{\theta} + \tilde{\phi}_k' \tilde{\theta}.
\]

Here, the unmodeled dynamics \( \tilde{\theta} \) impact the system output through the term \( \tilde{\phi}_k' \tilde{\theta} \). The unmodeled dynamics are characterized by certain saline features: (1) They are unknown but bounded when the system is stable. (2) They are better modeled as a deterministic uncertainty since they usually do not change randomly with time. (3) It is a multiplicative uncertainty,
in contrast to observation noise, which is commonly additive. While an additive uncertainty can be reduced by employing larger signals (namely a larger signal-to-noise ratio), a multiplicative uncertainty cannot be reduced by signal scaling. Consequently, its fundamental impact and reduction must be analyzed through different means. (4) The size of the unmodeled dynamics, \( \sum_{n=n_0}^{\infty} |a_n| \), is a monotonically decreasing function of the model complexity \( n_0 \).

2.5.2 System Uncertainties: Function Mismatch

When a system model involves a nonlinear function \( g(x) \), such as in the Wiener and Hammerstein models, the nonlinear function is often parameterized by \( g(x; \mu) \) with a finite parameter vector \( \mu \). This parameterization introduces model mismatch:

\[
\delta(x; \mu) = g(x) - g(x; \mu).
\]

Function mismatch is similar to multiplicative uncertainty, albeit in a nonlinear form, in which the reduction of \( \delta(x, \mu) \) cannot be achieved in general by signal scaling.

2.5.3 Sensor Bias and Drifts

A quantized sensor is characterized by its thresholds. In many applications, sensor thresholds may not be exactly known or change with time. System identification in which sensor thresholds are unknown must consider the thresholds as part of unknown parameters to be identified. Including thresholds in parameter vectors inevitably leads to a nonlinear structure.

2.5.4 Noise

The additive observation noise \( d_k \) in (2.1) may be modeled either as an unknown-but-bounded noise in a deterministic framework or as a random process in a stochastic framework.

In a deterministic framework, prior information on \( \{d_k\} \) is limited to its uniform bound \( |d_k| \leq \delta \). In other words, the uncertainty set is \( \Gamma_d = \{d_k \in \mathbb{R} : |d_k| \leq \delta\} \). Identification errors resulting from this type of observation noise are characterized by the worst-case bound over all possible noises in \( \Gamma_d \).

In contrast, in a stochastic framework, \( \{d_k\} \) is described as a random process. Typical cases include independent and identically distributed (i.i.d.) processes whose probability distribution function of \( d_1 \) is \( F(\cdot) \), and mixing processes for dependent noises.
2.5.5 Unknown Noise Characteristics

Identification methodologies and algorithms in this book utilize extensively information on noise distribution functions $F(\cdot)$ or density functions $f(\cdot)$. Such functions may not be available or may be subject to deviations and drifting. Unknown noise distribution functions compel their inclusion in identification. While it is possible to model distribution and density functions either parametrically or nonparametrically, the parameterization approach is more consistent with the methods of this book. As a result, in this book, $F(\cdot)$ is represented by a parameterized model $F(\cdot; \mu)$ when it is unknown.

2.5.6 Communication Channel Uncertainties

In networked systems, sensor outputs are not directly measured, but rather are transmitted through a communication channel, shown in Figure 2.3(b). When $s_k$ is transmitted through a communication channel, the received sequence $w_k$ is subject to channel noise and other uncertainties. When identification must be performed with observations on $w_k$, instead of $s_k$, channel noise and uncertainties will directly influence identification accuracy and convergence rates.

2.6 Notes

The basic system configurations presented in this chapter are representative in control systems, although they are not exhaustive. There are many other system settings that can be considered when essential issues are understood from the basic configurations. Quantization is an essential part of digital signals that have been extensively studied, mostly in uniformly spaced quantization. Its theoretical foundation and main properties have been studied in a different context beyond sampled data systems; see [1, 34, 80]. Quantized information processing occurs in many different applications beyond system identification [13, 38, 39, 73, 96]. Identification input design is an integral part of an identification experiment. This book is mostly limited to periodic signals due to their unique capability in providing input richness, in simplifying identification problems, and in their invariance when passing through systems. Most textbooks on system identification under stochastic formulations contain some discussions on uncertainties; see, for example, [17, 62]. The worst-case identification under set membership uncertainty is covered comprehensively in [66, 67, 68]. Models of noisy communication channels and their usage in information theory can be found in [22].
System Identification with Quantized Observations
Wang, L.Y.; Yin, G.; Zhang, J.-F.; Zhao, Y.
2010, XVIII, 317 p. 42 illus., Hardcover
A product of Birkhäuser Basel