Chapter 1

System Realization Theory for Linear Time-Invariant Systems

1.1 Introduction

Many algorithms have been established, some of them deterministic in nature, i.e. without considering noise in the measured data, and others stochastic, i.e. with formulations minimizing the noise uncertainty in the identification. During the 90s, building upon initial work by Gilbert and Kalman, several methods has been developed to identify most observable and controllable subspace of the system from given input-output (I/O) data [1, 2, 3, 4, 5, 6, 7]. Under the interaction of structure and control disciplines, the Eigensystem Realization Algorithm (ERA) [8] was developed for modal parameter identification and model reduction of dynamic systems using test data. The algorithm presents a unified framework for modal parameter identification based on the Markov parameters (i.e., pulse response) making it possible to construct a Hankel matrix as the basis for the realization of a discrete-time state-space model. A few years later at NASA, Juang developed a method for simultaneously identify a linear state-space model and the associated Kalman filter from noisy input-output measurements. Known as the Observer/Kalman Identification Algorithm (OKID) and formulated entirely in the time-domain, it computes the Markov parameters of a linear system, from which the state-space model and a corresponding observer are determined simultaneously [9]. The method relies on an observer equation to compress the dynamics of the system and efficiently estimate the associated system parameters (Markov parameters). In conjunction with the ERA, the method provides simultaneously both the Markov parameters and the Kalman gain, extracting all the possible information present in the data. The observer at the core of the method was proven to be the steady-state Kalman filter corresponding to the system to be identified. Later, the ERA with Data Correlation (ERA/DC) is developed [10, 11, 12, 13, 14] and while the ERA is, in essence, a least-squares fit to the pulse response measurements, the ERA/DC involves a fit to the output auto-correlation and cross-correlations over a defined number of lag values.

This document presents the fundamentals of linear time-invariant system identification and provides a common basis, definitions and notations to understand the techniques developed under the vast discipline of system identification. From discrete-time state-space models, controllability and observability to the famous OKID/ERA procedure, this document introduces the basic building blocks that are crucial notions in the field of system identification.

1.2 Time-Domain State-Space Models

1.2.1 Continuous-Time State-Space Models

The equations of motion for a finite-dimensional linear-dynamic system are a set of n first-order differential equations (Eq. (1.1a)) along with an initial condition $\boldsymbol{x}(t_0)$. The n-dimensional state $\boldsymbol{x}(t)$ is most often related to the output through the measurement equation Eq. (1.1b).

$$\dot{\boldsymbol{x}}(t) = A_c\boldsymbol{x}(t) + B_c\boldsymbol{u}(t), \tag{1.1a}$$

$$\boldsymbol{y}(t) = C\boldsymbol{x}(t) + D\boldsymbol{u}(t). \tag{1.1b}$$

The system of equations Eq. (1.1) constitutes a continuous-time state-space model of a dynamical system. Given the initial condition $\boldsymbol{x}(t_0)$ at some $t = t_0$, solving for $\boldsymbol{x}(t)$ for $t > t_0$ yields

$$\boldsymbol{x}(t) = e^{A_c(t-t_0)}\boldsymbol{x}(t_0) + \int_{t_0}^{t} e^{A_c(t-\tau)}B_c\boldsymbol{u}(\tau)\mathrm{d}\tau. \tag{1.2}$$

Without loss of generality, we will consider that $t_0 = 0$.

1.2.2 Discrete-Time State-Space Models

Dynamic systems are typically modeled by continuous-time or discrete-time equations. A close approximation of a continuous-time model can be obtained by a discrete one provided that the sampling rate is sufficiently high. A linear discrete system is most commonly described by an n^{th} order difference equation, the weighting sequence, or a discrete state-space model.
Let Δt be a constant time interval and $f = 1/\Delta t$ the sampling rate. Continuous versions of the A and B matrices are

$$A = e^{A_c\Delta t}, \tag{1.3a}$$

$$B = \int_0^{\Delta t} e^{A_c\tau}\mathrm{d}\tau B_c, \tag{1.3b}$$

$$\boldsymbol{x}(k+1) = \boldsymbol{x}((k+1)\Delta t), \tag{1.3c}$$

$$\boldsymbol{u}(k) = \boldsymbol{u}(k\Delta t). \tag{1.3d}$$

The discrete-time matrices A and B in Eqs (1.3a) and (1.3b) may be computed by the following series expansions:

$$A = e^{A_c\Delta t} = \sum_{i=0}^{\infty}\frac{1}{i!}\left[A_c\Delta t\right]^i, \tag{1.4a}$$

$$B = \int_0^{\Delta t} e^{A_c\tau}\mathrm{d}\tau B_c = \left[\sum_{i=0}^{\infty}\frac{1}{i!}A_c^i\left(\Delta t\right)^{i+1}\right]B_c. \tag{1.4b}$$

A sufficient condition for these series expansions to converge is that the continuous-time state matrix A_c is asymptotically stable in the sense that the real parts of all its eigenvalues are negative. If none of the eigenvalues of A_c are zero, the discrete-time matrix B may also be computed by

$$B = [A - I]\, A_c^{-1} B_c. \tag{1.5}$$

Therefore, a discrete-time invariant linear system can be represented by

$$\boldsymbol{x}(k+1) = A\boldsymbol{x}(k) + B\boldsymbol{u}(k) \tag{1.6a}$$
$$\boldsymbol{y}(k) = C\boldsymbol{x}(k) + D\boldsymbol{u}(k) \tag{1.6b}$$

together with an initial state vector $\boldsymbol{x}(0)$, where $\boldsymbol{x}$, $\boldsymbol{u}$ and $\boldsymbol{y}$ are the state, control input and output vectors respectively. The constant matrices A, B, C and D with appropriate dimensions represent the internal operation of the linear system, and are used to determine the system's response to any input.

1.2.3 Weighting Sequence Description and Markov Parameters

Solving for the state $\boldsymbol{x}(k)$ and the output $\boldsymbol{y}(k)$ with arbitrary initial condition $\boldsymbol{x}(0)$ in terms of the previous inputs $\boldsymbol{u}(i)$, $i = 0, 1, \dots, k$, yields

$$\boldsymbol{x}(k) = A^k\boldsymbol{x}(0) + \sum_{i=1}^{k} A^{i-1}B\boldsymbol{u}(k-i), \tag{1.7a}$$
$$\boldsymbol{y}(k) = CA^k\boldsymbol{x}(0) + \sum_{i=1}^{k} CA^{i-1}B\boldsymbol{u}(k-i) + D\boldsymbol{u}(k). \tag{1.7b}$$

It appears naturally that the constant matrices sequence

$$\begin{aligned} h_0 &= D,\\ h_1 &= CB,\\ h_2 &= CAB,\\ &\vdots\\ h_k &= CA^{k-1}B,\\ &\vdots \end{aligned} \tag{1.8}$$

plays an important role in identifying a mathematical model for linear dynamical systems. In fact, with zero initial condition $\boldsymbol{x}(0) = \boldsymbol{0}$, considering the response to a pulse sequence such that for $j = 1, 2, \dots, r$,

$$u_j(i) = \begin{cases} 1 & \text{for } i = 0 \\ 0 & \text{for } i = 1, 2, \dots \end{cases} \tag{1.9}$$

the r corresponding outputs $\left\{\boldsymbol{y}^{(j)}(i)\right\}_{i=1,2,\dots}$ can be assembled at each time step to recover the sequence $\{h_i\}_{i=1,2,\dots}$ as follows:

$$h_i = \begin{bmatrix}\boldsymbol{y}^{(1)}(i) & \boldsymbol{y}^{(2)}(i) & \cdots & \boldsymbol{y}^{(r)}(i)\end{bmatrix}. \tag{1.10}$$

The constant matrices in the sequence $\{h_i\}_{i=1,2,...}$ are known as system Markov parameters or, in short, *Markov parameters*. It is obvious that the matrices A, B, C, D are embedded in the Markov parameter sequence; undeniably, the determination of Markov parameters should be tantamount to system identification. The general form of the Markov parameters is thus

$$h_i = \begin{cases} D & i = 0, \\ CA^{i-1}B & i \geq 1, \\ 0 & i < 0. \end{cases} \tag{1.11}$$

1.3 Controllability and Observability

Before solving for the Markov parameters, it is of great importance to know whether all the states of a system can be controlled and/or observed since a solvable system of linear algebraic equations has a solution if and only if the rank of the system matrix is full. While controllability is concerned with whether one can design control input to steer the state to arbitrarily values, observability is concerned with whether without knowing the initial state, one can determine the state of a system given the input and the output.

1.3.1 Controllability in the discrete-time domain

A state $\boldsymbol{x}(q)$ is said to be *controllable* or *state-controllable* if this state can be reached from any initial state of the system in a finite time interval by some control action. If all states are controllable, the system is called *completely controllable* or simply *controllable*. Given A, B and $\boldsymbol{x}(0)$, the idea is to find the sufficient and necessary condition to determine how to reach $\boldsymbol{x}(q)$ without ambiguity. It is clear that since A and $\boldsymbol{x}(0)$ are given, it is therefore equivalent to determine $\boldsymbol{x}(q)$ or $\tilde{\boldsymbol{x}}(q) = \boldsymbol{x}(q) - A^q\boldsymbol{x}(0)$: to determine complete controllability, it is sufficient and necessary to determine whether the zero state (instead of all initial states) can be transferred to all final states.

The solution to the discrete representation at time $t = q\Delta t$ where Δt is the sampling period is

$$\boldsymbol{x}(q) = A^q\boldsymbol{x}(0) + \sum_{i=1}^{q} A^{i-1}B\boldsymbol{u}(q-i) \tag{1.12}$$

or in a compact matrix form

$$\boldsymbol{x}(q) = A^q\boldsymbol{x}(0) + \begin{bmatrix} B & AB & A^2B & \cdots & A^{q-1}B \end{bmatrix} \begin{bmatrix} \boldsymbol{u}(q-1) \\ \boldsymbol{u}(q-2) \\ \boldsymbol{u}(q-3) \\ \vdots \\ \boldsymbol{u}(0) \end{bmatrix}. \tag{1.13}$$

The expression of $\tilde{\boldsymbol{x}}(q)$ can be written as

$$\tilde{\boldsymbol{x}}(q) = \boldsymbol{R}^{(q)}\overline{\boldsymbol{u}} \tag{1.14}$$

where

$$\boldsymbol{R}^{(q)} = \begin{bmatrix} B & AB & A^2B & \cdots & A^{q-1}B \end{bmatrix}, \tag{1.15}$$

and

$$\overline{\boldsymbol{u}} = \begin{bmatrix} \boldsymbol{u}(q-1) \\ \boldsymbol{u}(q-2) \\ \boldsymbol{u}(q-3) \\ \vdots \\ \boldsymbol{u}(0) \end{bmatrix}. \tag{1.16}$$

$\boldsymbol{R}^{(q)}$ is called the controllability matrix. Reaching $\boldsymbol{x}(q)$ without ambiguity is thus equivalent to find a solution of Eq. (1.14) for $\overline{\boldsymbol{u}}$. Therefore, the discrete time-invariant linear system, Eq. 1.6, of order n is controllable if and only if the $n \times qr$ block controllability matrix $\boldsymbol{R}^{(q)}$ has rank n (assuming $qr > n$).

1.3.2 Observability in the discrete-time domain

A state $\boldsymbol{x}(p)$ is *observable* if the knowledge of the input $\boldsymbol{u}(k)$ and output $\boldsymbol{y}(k)$ over a finite time interval $0 \leq k \leq p-1$ completely determines the state $\boldsymbol{x}(p)$:

$$\boldsymbol{x}(p) = A^p\boldsymbol{x}(0) + \sum_{i=1}^{p} A^{i-1}B\boldsymbol{u}(p-i). \tag{1.17}$$

With knowledge of the system matrices A and B and the control input $\boldsymbol{u}(k)$, $0 \leq k \leq p-1$, it is necessary and sufficient to see whether the initial state $\boldsymbol{x}(0)$ can be completely determined from the output sequence $\boldsymbol{y}(k)$, $0 \leq k \leq p-1$. In a compact matrix form, we can write

$$\overline{\boldsymbol{y}} = \begin{bmatrix} \boldsymbol{y}(0) \\ \boldsymbol{y}(1) \\ \boldsymbol{y}(2) \\ \vdots \\ \boldsymbol{y}(p-1) \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{p-1} \end{bmatrix} \boldsymbol{x}(0) = \boldsymbol{O}^{(p)}\boldsymbol{x}(0), \tag{1.18}$$

where a unique solution for $\boldsymbol{x}(0)$ exists if and only if $\boldsymbol{O}^{(p)}$ has rank n (full rank, assuming $pm > n$). Thus, the discrete time-invariant linear system, Eq. 1.6, of order n is observable if and only if the $pm \times n$ block observability matrix $\boldsymbol{O}^{(p)}$ has rank n, where

$$\boldsymbol{O}^{(p)} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{p-1} \end{bmatrix}. \tag{1.19}$$

These two notions of controllability and observability will be of central attention in the next section for the development of the Eigensystem Realization Algorithm.

1.4 Coordinate Transformation

After having introduced the basic formulations of discrete-time invariant linear systems and before going in depth in the narrative of the Eigensystem Realization Algorithm, some results concerning coordinate transformation are introduced in this section.

In many problems, analysts need to use different coordinate systems in order to describe different quantities. A well-chosen coordinate system allows position and direction in space to be represented in a very convenient manner and allow sometimes to have some insights about the system itself. After all, two independent observers might well choose coordinate systems with different origins, and different orientations of the coordinate axes. The dynamic behavior of a mechanical system strongly depends upon its mathematical representation. This is due to the fact that nonlinearity is not an inherent attribute of a physical system, but rather dependent upon the mathematical description of the system's behavior. Ideally, one has an infinity of coordinate choices to represent the same physical system. In the study of celestial mechanics, the quest to find a judicious coordinate system led to the development of various canonical transformations. An extended discussion will be conducted in section concerning coordinate systems and transformations. This section only presents a few important results for linear discrete-time invariant systems.

Let a new state vector be defined such that

$$\tilde{\boldsymbol{x}}(k) = T\boldsymbol{x}(k), \tag{1.20}$$

where T is a nonsingular square matrix. Substitution of Eq. (1.20) into Eqs. (1.6) yields

$$\begin{cases} \tilde{\boldsymbol{x}}(k+1) = TAT^{-1}\tilde{\boldsymbol{x}}(k) + TB\boldsymbol{u}(k) \\ \boldsymbol{y}(k) \quad\quad = CT^{-1}\tilde{\boldsymbol{x}}(k) + \tilde{D}\boldsymbol{u}(k) \end{cases} \tag{1.21}$$

The effect of the input $\boldsymbol{u}(k)$ on the output $\boldsymbol{y}(k)$ is unchanged for the system. Thus the matrices $\{TAT^{-1}, TB, CT^{-1}, D\}$ describe the same input-output relationship as the matrices $\{A, B, C, D\}$:

$$\begin{cases} \boldsymbol{x}(k+1) = A\boldsymbol{x}(k) + B\boldsymbol{u}(k) \\ \boldsymbol{y}(k) \quad\quad = C\boldsymbol{x}(k) + D\boldsymbol{u}(k) \end{cases} \Leftrightarrow \begin{cases} \tilde{\boldsymbol{x}}(k+1) = \tilde{A}\tilde{\boldsymbol{x}}(k) + \tilde{B}\boldsymbol{u}(k) \\ \boldsymbol{y}(k) \quad\quad = \tilde{C}\tilde{\boldsymbol{x}}(k) + \tilde{D}\boldsymbol{u}(k) \end{cases} \tag{1.22}$$

with

$$\tilde{\boldsymbol{x}}(k) = T\boldsymbol{x}(k), \tag{1.23a}$$

$$\tilde{A} = TAT^{-1}, \tag{1.23b}$$

$$\tilde{B} = TB, \tag{1.23c}$$

$$\tilde{C} = CT^{-1}, \tag{1.23d}$$

$$\tilde{D} = D. \tag{1.23e}$$

This transformation is called a similarity transformation. The fact that the input-output relationship remains unchanged should also indicate that the pulse sequence, or Markov parameters, is invariant through coordinate transformation. Indeed, for $i \geq 1$,

$$\tilde{h}_i = \tilde{C}\tilde{A}^{i-1}\tilde{B} = CT^{-1}\left(TAT^{-1}\right)^{i-1}TB = CT^{-1}TA^{i-1}T^{-1}TB = CA^{i-1}B = h_i. \tag{1.24}$$

As a result, there are an infinite number of state-space representations that produce the same input-output description. Additionally, because matrices are similar if and only if they represent the same linear operator with respect to (possibly) different bases, similar matrices share all properties of their shared underlying operator (their rank in particular).

1.5 The Eigensystem Realization Algorithm (ERA)

The basic development of the state-space realization is attributed to Ho and Kalman [15] who introduced the important principles of minimum realization theory. The Ho-Kalman procedure uses the Hankel matrix to construct a state-space representation of a linear system from noise-free data. The methodology has been modified and substantially extended to develop the Eigensystem Realization Algorithm (ERA) [8] to identify modal parameters from noisy measurement data.

1.5.1 Hankel matrices

System realization begins by forming the generalized Hankel matrix composed of the Markov parameters:

$$\boldsymbol{H}_k^{(p,q)} = \begin{bmatrix} h_{k+1} & h_{k+2} & \cdots & h_{k+q} \\ h_{k+2} & h_{k+3} & \cdots & h_{k+q+1} \\ \vdots & \vdots & \ddots & \vdots \\ h_{k+p} & h_{k+p+1} & \cdots & h_{k+p+q-1} \end{bmatrix} = \boldsymbol{O}^{(p)} A^k \boldsymbol{R}^{(q)}. \tag{1.25}$$

For the case when $k = 0$,

$$\boldsymbol{H}_0^{(p,q)} = \begin{bmatrix} h_1 & h_2 & \cdots & h_q \\ h_2 & h_3 & \cdots & h_{q+1} \\ \vdots & \vdots & \ddots & \vdots \\ h_p & h_{p+1} & \cdots & h_{p+q-1} \end{bmatrix} = \boldsymbol{O}^{(p)} \boldsymbol{R}^{(q)}. \tag{1.26}$$

If $pm \geq n$ and $qr \geq n$, matrices $\boldsymbol{R}^{(q)}$ and $\boldsymbol{O}^{(p)}$ are of rank maximum n. If the system is controllable and observable, the block matrices $\boldsymbol{R}^{(q)}$ and $\boldsymbol{O}_p$ are of rank n. Therefore,

$$\text{rank}\left[\boldsymbol{H}_0^{(p,q)}\right] = \text{rank}\left[\boldsymbol{O}^{(p)}\boldsymbol{R}^{(q)}\right] \leq \min\left(\text{rank}\left[\boldsymbol{O}^{(p)}\right], \text{rank}\left[\boldsymbol{R}^{(q)}\right]\right) = n. \tag{1.27}$$

Since $\text{rank}\left[\boldsymbol{R}^{(q)}\right] = n$ ($\boldsymbol{R}^{(q)}$ is non-singular, the system is assumed to be controllable), multiplying both sides by $\boldsymbol{R}^{(q)^\dagger}$ yields

$$n = \text{rank}\left[\boldsymbol{O}^{(p)}\right] = \text{rank}\left[\left(\boldsymbol{O}^{(p)}\boldsymbol{R}^{(q)}\right)\boldsymbol{R}^{(q)^\dagger}\right] \leq \text{rank}\left[\boldsymbol{O}^{(p)}\boldsymbol{R}^{(q)}\right] = \text{rank}\left[\boldsymbol{H}_0^{(p,q)}\right]. \tag{1.28}$$

Hence we have

$$\text{rank}\left[\boldsymbol{H}_0^{(p,q)}\right] = n. \tag{1.29}$$

If the order is n, then the minimum dimension of the state matrix A is $n \times n$ and therefore, for any $k \geq 0$,

$$\text{rank}\left[\boldsymbol{H}_k^{(p,q)}\right] = n. \tag{1.30}$$

Thus, it appears that identifying the number of dominant singular values of the Hankel matrix provides an indication about the unknown order of the reduced model to be identified.

1.5.2 Hankel Norm Approximation

As described in the previous section, a singular value decomposition on the Hankel matrix provides an insight about the order of the system. Even if more advanced methods for distinguishing true modes from noise modes exist, a simple singular value plot often allows the engineer to determine the order of the system. Thus, it is possible to observe the following approximation

$$\boldsymbol{H}_0^{(p,q)} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\intercal = \begin{bmatrix}\boldsymbol{U}^{(n)} & \boldsymbol{U}^{(0)}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Sigma}^{(n)} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{\Sigma}^{(0)}\end{bmatrix}\begin{bmatrix}\boldsymbol{V}^{(n)\intercal}\\ \boldsymbol{V}^{(0)\intercal}\end{bmatrix} \tag{1.31}$$

$$= \boldsymbol{U}^{(n)}\boldsymbol{\Sigma}^{(n)}\boldsymbol{V}^{(n)\intercal} + \underbrace{\boldsymbol{U}^{(0)}\boldsymbol{\Sigma}^{(0)}\boldsymbol{V}^{(0)\intercal}}_{\simeq\boldsymbol{0}} \tag{1.32}$$

$$\simeq \boldsymbol{U}^{(n)}\boldsymbol{\Sigma}^{(n)}\boldsymbol{V}^{(n)\intercal} \tag{1.33}$$

where $\boldsymbol{U}^{(n)}$ and $\boldsymbol{V}^{(n)}$ are orthonormal matrices:

$$\boldsymbol{U}^{(n)\intercal}\boldsymbol{U}^{(n)} = \boldsymbol{V}^{(n)\intercal}\boldsymbol{V}^{(n)} = \boldsymbol{I}^{(n)}. \tag{1.34}$$

Since $\boldsymbol{H}_0^{(p,q)}$ is primarily represented by the controllability and observability matrices, a balanced factorization leads to

$$\boldsymbol{H}_0^{(p,q)} = \boldsymbol{U}^{(n)}\boldsymbol{\Sigma}^{(n)}\boldsymbol{V}^{(n)\intercal} = \boldsymbol{O}^{(p)}\boldsymbol{R}^{(q)} \Rightarrow \begin{cases}\boldsymbol{O}^{(p)} = \boldsymbol{U}^{(n)}\boldsymbol{\Sigma}^{(n)1/2}\\ \boldsymbol{R}^{(q)} = \boldsymbol{\Sigma}^{(n)1/2}\boldsymbol{V}^{(n)\intercal}\end{cases}. \tag{1.35}$$

This choice makes both $\boldsymbol{O}^{(p)}$ and $\boldsymbol{R}^{(q)}$ balanced. Notice that $\boldsymbol{R}^{(q)}\boldsymbol{R}^{(q)\intercal} = \boldsymbol{O}^{(p)\intercal}\boldsymbol{O}^{(p)} = \boldsymbol{\Sigma}^{(n)}$. The fact that the controllability and observability matrices are equal and diagonal implies that the realized system is as controllable as it is observable. This property is called an internally balanced realization. It means that the signal transfer from the input to the state and then from the state to the output are similar and balanced.

1.5.3 Minimum Realization

With $k = 1$ in Eq. (1.25), one obtains that

$$\boldsymbol{H}_1^{(p,q)} = \boldsymbol{O}^{(p)}A\boldsymbol{R}^{(q)} = \boldsymbol{U}^{(n)}\boldsymbol{\Sigma}^{(n)1/2}A\boldsymbol{\Sigma}^{(n)1/2}\boldsymbol{V}^{(n)\intercal}, \tag{1.36}$$

and a solution for the state matrix A becomes

$$\hat{A} = \boldsymbol{O}^{(p)\dagger}\boldsymbol{H}_1^{(p,q)}\boldsymbol{R}^{(q)\dagger} = \boldsymbol{\Sigma}^{(n)-1/2}\boldsymbol{U}^{(n)\intercal}\boldsymbol{H}_1^{(p,q)}\boldsymbol{V}^{(n)}\boldsymbol{\Sigma}^{(n)-1/2}. \tag{1.37}$$

Moreover, from Eq (1.15) and (1.19), it is clear that the first r columns of $\boldsymbol{R}^{(q)}$ form the input matrix B whereas the first m rows of $\boldsymbol{O}^{(p)}$ form the output matrix C. Defining O_i as a null matrix of order i, I_i as an identity matrix of order i and

$$\boldsymbol{E}^{(m)\intercal} = \begin{bmatrix}I_m & O_m & \cdots & O_m\end{bmatrix}, \tag{1.38a}$$

$$\boldsymbol{E}^{(r)\intercal} = \begin{bmatrix}I_r & O_r & \cdots & O_r\end{bmatrix}, \tag{1.38b}$$

a minimum realization is given by

$$\hat{A} = \boldsymbol{O}^{(p)\dagger}\boldsymbol{H}_1^{(p,q)}\boldsymbol{R}^{(q)\dagger} = \boldsymbol{\Sigma}^{(n)^{-1/2}}\boldsymbol{U}^{(n)\intercal}\boldsymbol{H}_1^{(p,q)}\boldsymbol{V}^{(n)}\boldsymbol{\Sigma}^{(n)^{-1/2}}, \tag{1.39a}$$

$$\hat{B} = \boldsymbol{R}^{(q)}\boldsymbol{E}^{(r)} = \boldsymbol{\Sigma}^{(n)^{1/2}}\boldsymbol{V}^{(n)\intercal}\boldsymbol{E}^{(r)}, \tag{1.39b}$$

$$\hat{C} = \boldsymbol{E}^{(m)\intercal}\boldsymbol{O}^{(p)} = \boldsymbol{E}^{(m)\intercal}\boldsymbol{U}^{(n)}\boldsymbol{\Sigma}^{(n)^{1/2}}, \tag{1.39c}$$

$$\hat{D} = h_0. \tag{1.39d}$$

The realized discrete-time model represented by the matrices $\hat{A}$, $\hat{B}$, $\hat{C}$ and $\hat{D}$ can be transformed to the continuous-time model. The system frequencies and damping may then be computed from the eigenvalues of the estimated continuous-time state matrix. The eigenvectors allow a transformation of the realization to modal space and hence a determination of the complex (or damped) mode shapes and the initial modal amplitudes (or modal participation factors).

1.6 Observer Kalman Filter Identification

Most techniques to identify the Markov parameters sequence Eq. (1.11) are based on the Fast Fourier Transform (FFT) of the inputs and measured outputs to compute the Frequency Response Functions (FRFs), and then use the Inverse Discrete Fourier Transform (IDFT) to compute the sampled pulse response histories. The discrete nature of the FFT causes one to obtain pulse response rather than impulse response, and a somewhat rich input is required to prevent numerical ill-conditioning. Indeed, the FRF is a ratio between the output and input DFT transform coefficients which requires the input signal to be rich in frequencies so that the corresponding quantity is invertible. However, considerable information can be deduced simply by observing frequency response functions, justifying why FRFs are still generated so often. Another approach is to solve directly in the time domain for the system Markov parameters from the input and output data. In [9], a method has been developed to compute the Markov parameters of a linear system in the time-domain. A drawback of this direct time-domain method is the need to invert an input matrix which necessarily becomes particularly large for lightly damped systems. Rather than identifying the system Markov parameters which may exhibit very slow decay, one can use an asymptotically stable observer to form a stable discrete state-space model for the system to be identified. The method is referred as the Observer/Kalman filter Identification algorithm (OKID) and is a procedure where the state-space model and a corresponding observer are determined simultaneously [16, 17].

1.6.1 Classical Formulation

Considering a sequence of l elements, assuming zero initial condition $\boldsymbol{x}(0) = \boldsymbol{0}$:

$$\boldsymbol{y}(l-1) = \sum_{i=1}^{l-1} CA^{i-1}B\boldsymbol{u}(l-1-i) + D\boldsymbol{u}(l-1). \tag{1.40}$$

In a matrix form, Eq. (1.40) is written as

$$\bar{\boldsymbol{y}} = \boldsymbol{Y}\boldsymbol{U} \tag{1.41}$$

with

$$\bar{\boldsymbol{y}} = \begin{bmatrix} \boldsymbol{y}(0) & \boldsymbol{y}(1) & \boldsymbol{y}(2) & \cdots & \boldsymbol{y}(l-1) \end{bmatrix}, \tag{1.42a}$$

$$\boldsymbol{Y} = \begin{bmatrix} D & CB & CAB & \cdots & CA^{l-2}B \end{bmatrix}, \tag{1.42b}$$

$$\boldsymbol{U} = \begin{bmatrix} \boldsymbol{u}(0) & \boldsymbol{u}(1) & \boldsymbol{u}(2) & \cdots & \boldsymbol{u}(l-1) \\ & \boldsymbol{u}(0) & \boldsymbol{u}(1) & \cdots & \boldsymbol{u}(l-2) \\ & & \boldsymbol{u}(0) & \cdots & \boldsymbol{u}(l-3) \\ & & & \ddots & \vdots \\ & & & & \boldsymbol{u}(0) \end{bmatrix}. \tag{1.42c}$$

The matrix $\bar{\boldsymbol{y}}$ is an $m \times l$ output data matrix and the matrix $\boldsymbol{Y}$, of dimension $m \times rl$, contains all the Markov parameters to be determined. As summarized in Table 1.1, Eq. (1.41) is insolvable in the multi-input multi-output case in general: the solution $\boldsymbol{Y}$ is not unique whereas it should be (Markov parameters must be unique for a finite-dimensional linear system). The matrix $\boldsymbol{Y}$ can only be uniquely determined from this set of equations for $r = 1$. However, even in this case, if the input has zero initial value or is not rich enough in frequency content or if anything makes the matrix $\boldsymbol{U}$ ill-conditioned, the matrix $\boldsymbol{Y} = \bar{\boldsymbol{y}}\boldsymbol{U}^{-1}$ cannot be accurately computed.

Table 1.1: Equations vs Unknowns for Eq. (1.41)

# Equations	# Unknowns
$m \times l$	$m \times rl$

However, if A is asymptotically stable so that for some l_0, $CA^kB \simeq 0$ for all time steps $k \geq l_0$, Eq. (1.41) can be approximated by

$$\bar{\boldsymbol{y}} \simeq \widetilde{\boldsymbol{Y}}\widetilde{\boldsymbol{U}}, \tag{1.43}$$

with

$$\bar{\boldsymbol{y}} = \begin{bmatrix} \boldsymbol{y}(0) & \boldsymbol{y}(1) & \boldsymbol{y}(2) & \cdots & \boldsymbol{y}(l-1) \end{bmatrix}, \tag{1.44a}$$

$$\widetilde{\boldsymbol{Y}} = \begin{bmatrix} D & CB & CAB & \cdots & CA^{l_0-1}B \end{bmatrix}, \tag{1.44b}$$

$$\widetilde{\boldsymbol{U}} = \begin{bmatrix} \boldsymbol{u}(0) & \boldsymbol{u}(1) & \boldsymbol{u}(2) & \cdots & \boldsymbol{u}(l_0) & \cdots & \boldsymbol{u}(l-1) \\ & \boldsymbol{u}(0) & \boldsymbol{u}(1) & \cdots & \boldsymbol{u}(l_0-1) & \cdots & \boldsymbol{u}(l-2) \\ & & \boldsymbol{u}(0) & \cdots & \boldsymbol{u}(l_0-2) & \cdots & \boldsymbol{u}(l-3) \\ & & & \ddots & \vdots & \cdots & \vdots \\ & & & & \boldsymbol{u}(0) & \cdots & \boldsymbol{u}(l-l_0-1) \end{bmatrix}. \tag{1.44c}$$

Choose the data record length l greater than $r(l_0+1)$ and Eq. (1.43) indicates that there are more equations than constraints, as summarized in Table 1.2. It follows that the first l_0+1 Markov parameters approximately satisfy

$$\widetilde{\boldsymbol{Y}} = \bar{\boldsymbol{y}}\widetilde{\boldsymbol{U}}^{\dagger}, \tag{1.45}$$

and the approximation error decreases as l_0 increases.

Table 1.2: Equations vs Unknowns for Eq. (1.43)

# Equations	# Unknowns
$m \times l$	$m \times r(l_0 + 1)$

Unfortunately, for lightly damped structures, the integer l_0 and thus the the data length l required to make the approximation in Eq. (1.43) valid becomes impractically large in the sense that the size of the matrix $\widetilde{\boldsymbol{U}}$ is too large to solve for its pseudo-inverse numerically. A solution to artificially increase the damping of the system in order to allow a decent numerical solution is to add a feedback loop to make the system as stable as desired.

1.6.2 State-space Observer Model

In practice, the primary purpose of introducing an observer is an artifice to compress the data and improve system identification results. A state estimator, also known as an observer, can be used to provide an estimate of the system state from input and output measurements. Add and subtract the term $G\boldsymbol{y}(k)$ to the right-hand side of the state equation in Eq. (1.6a) to yield

$$\begin{aligned} \boldsymbol{x}(k+1) &= A\boldsymbol{x}(k) + B\boldsymbol{u}(k) + G\boldsymbol{y}(k) - G\boldsymbol{y}(k) && (1.46a) \\ &= (A + GC)\boldsymbol{x}(k) + (B + GD)\boldsymbol{u}(k) - G\boldsymbol{y}(k) && (1.46b) \\ &= \bar{A}\boldsymbol{x}(k) + \bar{B}\boldsymbol{v}(k) && (1.46c) \end{aligned}$$

where

$$\begin{aligned} \bar{A} &= A + GC, && (1.47a) \\ \bar{B} &= \begin{bmatrix} B + GD & -G \end{bmatrix}, && (1.47b) \\ \boldsymbol{v}(k) &= \begin{bmatrix} \boldsymbol{u}(k) \\ \boldsymbol{y}(k) \end{bmatrix}, && (1.47c) \end{aligned}$$

and G is an arbitrary matrix chosen to make the matrix $\bar{A}$ as stable as desired. The Markov parameters of this observer system are referred as *observer Markov parameters*. The new input-output in matrix form is therefore

$$\bar{\boldsymbol{y}} = \bar{\boldsymbol{Y}}\boldsymbol{V} \tag{1.48}$$

with

$$\bar{\boldsymbol{y}} = \begin{bmatrix} \boldsymbol{y}(0) & \boldsymbol{y}(1) & \boldsymbol{y}(2) & \cdots & \boldsymbol{y}(l-1) \end{bmatrix}, \tag{1.49}$$

$$\bar{\boldsymbol{Y}} = \begin{bmatrix} D & C\bar{B} & C\bar{A}\bar{B} & \cdots & C\bar{A}^{l-2}\bar{B} \end{bmatrix}, \tag{1.50}$$

$$\boldsymbol{V} = \begin{bmatrix} \boldsymbol{u}(0) & \boldsymbol{u}(1) & \boldsymbol{u}(2) & \cdots & \boldsymbol{u}(l-1) \\ & \boldsymbol{v}(0) & \boldsymbol{v}(1) & \cdots & \boldsymbol{v}(l-2) \\ & & \boldsymbol{v}(0) & \cdots & \boldsymbol{v}(l-3) \\ & & & \ddots & \vdots \\ & & & & \boldsymbol{v}(0) \end{bmatrix}. \tag{1.51}$$

Table 1.3: Equations vs Unknowns for Eq. (1.48)

# Equations	# Unknowns
$m \times l$	$m \times ((r+m)(l-1)+r)$

Similarly as before, when $C\bar{A}^k\bar{B} \simeq 0$ for all time steps $k \geq l_0$ for some l_0, Eq. (1.48) can be approximated by

$$\bar{\boldsymbol{y}} \simeq \widetilde{\bar{\boldsymbol{Y}}}\widetilde{\boldsymbol{V}}, \tag{1.52}$$

with

$$\bar{\boldsymbol{y}} = \begin{bmatrix}\boldsymbol{y}(0) & \boldsymbol{y}(1) & \boldsymbol{y}(2) & \cdots & \boldsymbol{y}(l-1)\end{bmatrix}, \tag{1.53}$$

$$\widetilde{\bar{\boldsymbol{Y}}} = \begin{bmatrix}D & C\bar{B} & C\bar{A}\bar{B} & \cdots & C\bar{A}^{l_0-1}\bar{B}\end{bmatrix}, \tag{1.54}$$

$$\widetilde{\boldsymbol{V}} = \begin{bmatrix}\boldsymbol{u}(0) & \boldsymbol{u}(1) & \boldsymbol{u}(2) & \cdots & \boldsymbol{u}(l_0) & \cdots & \boldsymbol{u}(l-1) \\ & \boldsymbol{v}(0) & \boldsymbol{v}(1) & \cdots & \boldsymbol{v}(l_0-1) & \cdots & \boldsymbol{v}(l-2) \\ & & \boldsymbol{v}(0) & \cdots & \boldsymbol{v}(l_0-2) & \cdots & \boldsymbol{v}(l-3) \\ & & & \ddots & \vdots & \cdots & \vdots \\ & & & & \boldsymbol{v}(0) & \cdots & \boldsymbol{v}(l-l_0-1)\end{bmatrix}. \tag{1.55}$$

Table 1.4: Equations vs Unknowns for Eq. (1.52)

# Equations	# Unknowns
$m \times l$	$m \times ((r+m)l_0+r)$

The first l_0+1 observer Markov parameters approximately satisfy

$$\widetilde{\bar{\boldsymbol{Y}}} = \bar{\boldsymbol{y}}\widetilde{\boldsymbol{V}}^\dagger, \tag{1.56}$$

and the approximation error decreases as l_0 increases. To solve for $\widetilde{\bar{\boldsymbol{Y}}}$ uniquely, all the rows of $\widetilde{\boldsymbol{V}}$ must be linearly independent. Furthermore, to minimize any numerical error due to the computation of the pseudo-inverse, the rows of $\widetilde{\boldsymbol{V}}$ should be chosen as independently as possible. As a result, the maximum value of l_0 is the number that maximizes the quantity $(r+m)l_0+r \leq l$ of independent rows of $\widetilde{\boldsymbol{V}}$. The maximum l_0 means the upper bound of the order of the deadbeat observer. Furthermore, it is known that the rank of a sufficiently large Hankel matrix $\boldsymbol{H}_0^{(p,q)}$ is the order of the controllable and observable part of the system (the identified state matrix $\hat{A}$ represents only the controllable and observable part of the system). The size of the Hankel matrix is $pm \times qr$ comprised of $p+q-1$ Markov parameters; with $p=q$, this means $2p-1$ Markov parameters. If l_0 is the number of Markov parameters calculated through OKID, it means that $l_0 = 2p-1$. Assuming $qr > n$, the maximum rank of $\boldsymbol{H}_0^{(p,q)}$ is thus mp. If p is chosen such that $mp \geq n$, then a realized state matrix $\hat{A}$ with order n should exist. Therefore, the number of Markov parameters computed, l_0, must be chosen such that

$$mp \geq n \Leftrightarrow m\frac{l_0+1}{2} \geq n, \tag{1.57}$$

where m is the number of outputs and n is the order of the system. To conclude, the lower and upper bounds on l_0 are

$$\frac{2n}{m} - 1 \leq l_0 \leq \frac{l-r}{r+m} \tag{1.58}$$

with l being the length of the input signal considered.

The observer Markov parameters can then be used to compute the Markov parameters and the matrices A, B, C and D.

1.6.3 Computation of Markov parameters from observer Markov parameters

To recover the system Markov parameters from the observer Markov parameters, write

$$\bar{h}_0 = D, \tag{1.59a}$$

$$\bar{h}_k = C\bar{A}^{k-1}\bar{B} \tag{1.59b}$$

$$= \begin{bmatrix} C(A+GC)^{k-1}(B+GD) & -C(A+GC)^{k-1}G \end{bmatrix} \tag{1.59c}$$

$$= \begin{bmatrix} \bar{h}_k^{(1)} & -\bar{h}_k^{(2)} \end{bmatrix} \tag{1.59d}$$

Thus, the Markov parameter h_1 of the system is simply

$$h_1 = CB = C(B+GD) - (CG)D = \bar{h}_1^{(1)} - \bar{h}_1^{(2)}D = \bar{h}_1^{(1)} - \bar{h}_1^{(2)}h_0. \tag{1.60}$$

Considering the product

$$\bar{h}_2^{(1)} = C(A+GC)(B+GD) = CAB + CGCB + C(A+GC)GD = h_2 + \bar{h}_1^{(2)}h_1 + \bar{h}_2^{(2)}h_0, \tag{1.61}$$

the next Markov parameter is

$$h_2 = CAB = \bar{h}_2^{(1)} - \bar{h}_1^{(2)}h_1 - \bar{h}_2^{(2)}h_0. \tag{1.62}$$

Note that the sum of subscript(s) of each individual term both sides is identical. By induction, the general relationship between the actual system Markov parameters and the observer Markov parameters is

$$h_0 = \bar{h}_0, \tag{1.63a}$$

$$h_k = \bar{h}_k^{(1)} - \sum_{i=1}^{k} \bar{h}_i^{(2)} h_{k-i}, \quad \text{for } k \geq 1. \tag{1.63b}$$

1.7 The ERA from initial condition response

The state-variable response of a system described by Eq. (1.6) with zero input and an arbitrary set of initial conditions $\boldsymbol{x}(0)$ is:

$$\boldsymbol{x}(k) = A^k \boldsymbol{x}(0), \tag{1.64a}$$

$$\boldsymbol{y}(k) = CA^k \boldsymbol{x}(0). \tag{1.64b}$$

In that situation, the significance of previously defined Markov parameters is gone. As they are originally defined as pulse response, there is no worthwile definition for these matrices here. Similarly, there is no meaning for controllability in this case as the input control is set to zero. However, referring to section 2.2.2, the concept of observability is still relevant. Even though observability and controllability of a linear system are mathematical duals, the concept of observability is just a measure of how well internal states of a system can be inferred from knowledge of its external outputs. The same way observability was previously defined, the discrete time-invariant linear system, Eq. (1.64), of order n is observable if and only if the $pm \times n$ block observability matrix $\boldsymbol{O}^{(p)}$ has rank n, where

$$\boldsymbol{O}^{(p)} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{p-1} \end{bmatrix}. \tag{1.65}$$

Even though controllability has no substantial meaning, it is possible to define a controllability-like matrix, named $\boldsymbol{Q}^{(q)}$, that gathers the state variable at different times:

$$\boldsymbol{Q}^{(q)} = \begin{bmatrix} \boldsymbol{x}(0) & A\boldsymbol{x}(0) & A^2\boldsymbol{x}(0) & \cdots & A^{q-1}\boldsymbol{x}(0) \end{bmatrix}. \tag{1.66}$$

Since $\boldsymbol{x}(0)$ is a n-dimensional vector and $A \in \mathbb{R}^{n\times n}$, $\boldsymbol{Q}^{(q)} \in \mathbb{R}^{n\times q}$ and has rank n from the moment $q \geq n$.

Let's now define a Hankel matrix as

$$\boldsymbol{H}_k^{(p,q)} = \begin{bmatrix} \boldsymbol{y}(k) & \boldsymbol{y}(k+1) & \cdots & \boldsymbol{y}(k+q-1) \\ \boldsymbol{y}(k+1) & \boldsymbol{y}(k+2) & \cdots & \boldsymbol{y}(k+q) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{y}(k+p-1) & \boldsymbol{y}(k+p) & \cdots & \boldsymbol{y}(k+p+q-2) \end{bmatrix} = \boldsymbol{O}^{(p)} A^k \boldsymbol{Q}^{(q)}. \tag{1.67}$$

For the case when $k = 0$,

$$\boldsymbol{H}_0^{(p,q)} = \begin{bmatrix} \boldsymbol{y}(0) & \boldsymbol{y}(1) & \cdots & \boldsymbol{y}(q-1) \\ \boldsymbol{y}(1) & \boldsymbol{y}(2) & \cdots & \boldsymbol{y}(q) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{y}(p-1) & \boldsymbol{y}(p) & \cdots & \boldsymbol{y}(p+q-2) \end{bmatrix} = \boldsymbol{O}^{(p)} \boldsymbol{Q}^{(q)}. \tag{1.68}$$

If $pm \geq n$ and $q \geq n$, matrices $\boldsymbol{Q}^{(q)}$ and $\boldsymbol{O}^{(p)}$ are of rank maximum n. If the system is observable, the block matrix $\boldsymbol{O}_p$ is of rank n and

$$\text{rank}\left[\boldsymbol{H}_k^{(p,q)}\right] = n. \tag{1.69}$$

Following the exact same steps as for the classical ERA, it leads to

$$\boldsymbol{H}_0^{(p,q)} = \boldsymbol{U}^{(n)}\boldsymbol{\Sigma}^{(n)}\boldsymbol{V}^{(n)\intercal} = \boldsymbol{O}^{(p)}\boldsymbol{Q}^{(q)} \Rightarrow \begin{cases} \boldsymbol{O}^{(p)} = \boldsymbol{U}^{(n)}{\boldsymbol{\Sigma}^{(n)}}^{1/2} \\ \boldsymbol{Q}^{(q)} = {\boldsymbol{\Sigma}^{(n)}}^{1/2}\boldsymbol{V}^{(n)\intercal} \end{cases}, \tag{1.70}$$

and a minimum realization is given by

$$\hat{A} = \boldsymbol{O}^{(p)\dagger}\boldsymbol{H}_1^{(p,q)}\boldsymbol{Q}^{(q)\dagger} = \boldsymbol{\Sigma}^{(n)^{-1/2}}\boldsymbol{U}^{(n)\intercal}\boldsymbol{H}_1^{(p,q)}\boldsymbol{V}^{(n)}\boldsymbol{\Sigma}^{(n)^{-1/2}}, \tag{1.71a}$$

$$\hat{C} = \boldsymbol{E}^{(m)\intercal}\boldsymbol{O}^{(p)} = \boldsymbol{E}^{(m)\intercal}\boldsymbol{U}^{(n)}\boldsymbol{\Sigma}^{(n)^{1/2}}, \tag{1.71b}$$

$$\hat{\boldsymbol{x}}_0 = \boldsymbol{Q}^{(q)}\boldsymbol{E}^{(1)} = \boldsymbol{\Sigma}^{(n)^{1/2}}\boldsymbol{V}^{(n)\intercal}\boldsymbol{E}^{(1)}. \tag{1.71c}$$

Note that this formulation is very close to the classical ERA formulation.

Table 1.5: Classical ERA vs ERA with Initial Condition Response

Classical ERA	ERA with Initial Condition Response
$\boldsymbol{H}_k^{(p,q)} \in \mathbb{R}^{pm\times qr}$	$\boldsymbol{H}_k^{(p,q)} \in \mathbb{R}^{pm\times q}$
$\boldsymbol{R}^{(q)} = \begin{bmatrix} B & AB & \cdots & A^{q-1}B \end{bmatrix} \in \mathbb{R}^{n\times qr}$	$\boldsymbol{Q}^{(q)} = \begin{bmatrix} \boldsymbol{x}(0) & A\boldsymbol{x}(0) & \cdots & A^{q-1}\boldsymbol{x}(0) \end{bmatrix} \in \mathbb{R}^{n\times q}$
$\boldsymbol{V}^{(n)\intercal} \in \mathbb{R}^{n\times qr}$	$\boldsymbol{V}^{(n)\intercal} \in \mathbb{R}^{n\times q}$
$\hat{B} = \boldsymbol{R}^{(q)}\boldsymbol{E}^{(r)}$	$\hat{\boldsymbol{x}}_0 = \boldsymbol{Q}^{(q)}\boldsymbol{E}^{(1)}$

1.8 The ERA with Data Correlations (ERA/DC)

The Eigensystem Realization Algorithm with Data Correlations (ERA/DC) includes an additional fit to output correlations whereas the ERA is basically a least-square fit to the pulse response measurements only. The bias terms affecting the ERA when noise is present can, in principle, be omitted in the ERA/DC by properly tuning some of the parameters. The computational steps of the ERA/DC are outlined in this section.

1.8.1 Block Correlation Hankel Matrices

The ERA method with Data Correlation (ERA/DC) requires the definition of a square matrix of order $\gamma = mp$,

$$\boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(k) = \boldsymbol{H}(k)\boldsymbol{H}(0)^\intercal \tag{1.72}$$

$$= \begin{bmatrix} h_{k+1} & h_{k+2} & \cdots & h_{k+q} \\ h_{k+2} & h_{k+3} & \cdots & h_{k+q+1} \\ \vdots & \vdots & \ddots & \vdots \\ h_{k+p} & h_{k+p+1} & \cdots & h_{k+p+q-1} \end{bmatrix} \begin{bmatrix} h_1 & h_2 & \cdots & h_q \\ h_2 & h_3 & \cdots & h_{q+1} \\ \vdots & \vdots & \ddots & \vdots \\ h_p & h_{p+1} & \cdots & h_{p+q-1} \end{bmatrix}^\intercal \tag{1.73}$$

$$= \begin{bmatrix} \sum_{i=1}^{q} h_{k+i}h_i^\intercal & \sum_{i=1}^{q} h_{k+i}h_{i+1}^\intercal & \cdots & \sum_{i=1}^{q} h_{k+i}h_{p+i-1}^\intercal \\ \sum_{i=1}^{q} h_{k+i+1}h_i^\intercal & \sum_{i=1}^{q} h_{k+i+1}h_{i+1}^\intercal & \cdots & \sum_{i=1}^{q} h_{k+i+1}h_{p+i-1}^\intercal \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{q} h_{k+p+i-1}h_i^\intercal & \sum_{i=1}^{q} h_{k+p+i-1}h_{i+1}^\intercal & \cdots & \sum_{i=1}^{q} h_{k+p+i-1}h_{p+i-1}^\intercal \end{bmatrix} \tag{1.74}$$

Note that the data correlation matrix $\boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(k)$ can be smaller in size than the Hankel matrix $\boldsymbol{H}(k)$ if $qr \leq pm$.
For the case when $k = 0$, the correlation matrix $\boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(0)$ becomes

$$\boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(0) = \boldsymbol{H}(0)\boldsymbol{H}(0)^{\intercal} \tag{1.75}$$

$$= \begin{bmatrix} h_1 & h_2 & \cdots & h_q \\ h_2 & h_3 & \cdots & h_{q+1} \\ \vdots & \vdots & \ddots & \vdots \\ h_p & h_{p+1} & \cdots & h_{p+q-1} \end{bmatrix} \begin{bmatrix} h_1 & h_2 & \cdots & h_q \\ h_2 & h_3 & \cdots & h_{q+1} \\ \vdots & \vdots & \ddots & \vdots \\ h_p & h_{p+1} & \cdots & h_{p+q-1} \end{bmatrix}^{\intercal} \tag{1.76}$$

$$= \begin{bmatrix} \sum_{i=1}^{q} h_i h_i^{\intercal} & \sum_{i=1}^{q} h_i h_{i+1}^{\intercal} & \cdots & \sum_{i=1}^{q} h_i h_{p+i-1}^{\intercal} \\ \sum_{i=1}^{q} h_{i+1} h_i^{\intercal} & \sum_{i=1}^{q} h_{i+1} h_{i+1}^{\intercal} & \cdots & \sum_{i=1}^{q} h_{i+1} h_{p+i-1}^{\intercal} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{q} h_{p+i-1} h_i^{\intercal} & \sum_{i=1}^{q} h_{p+i-1} h_{i+1}^{\intercal} & \cdots & \sum_{i=1}^{q} h_{p+i-1} h_{p+i-1}^{\intercal} \end{bmatrix} \tag{1.77}$$

The matrix $\boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(0)$ consists of auto-correlations of Markov parameters such as $\sum_{i=1}^{q} h_i h_i^{\intercal}$ and cross-correlations between outputs such as $\sum_{i=1}^{q} h_i h_{i+1}^{\intercal}$ at lag time values in the range $\pm p$, summed over q values. If noises in the Markov parameters are not correlated, the correlation matrix $\boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(0)$ will contain less noise than the Hankel matrix $\boldsymbol{H}(0)$.

In terms of controllability and observability matrices, $\boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(k)$ can be written as

$$\boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(k) = \boldsymbol{O}_p A^k \boldsymbol{R}_q \boldsymbol{R}_q^{\intercal} \boldsymbol{O}_p^{\intercal} = \boldsymbol{O}_p A^k \boldsymbol{R}_{\gamma}, \tag{1.78}$$

where $\boldsymbol{R}_{\gamma} = \boldsymbol{R}_q \boldsymbol{R}_q^{\intercal} \boldsymbol{O}_p^{\intercal}$.
The data correlation matrix $\boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(k)$ can be used to form a block correlation Hankel matrix

$$\begin{aligned} \boldsymbol{\mathcal{H}}(k) &= \begin{bmatrix} \boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(k) & \boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(k+\tau) & \cdots & \boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(k+\zeta\tau) \\ \boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(k+\tau) & \boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(k+2\tau) & \cdots & \boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(k+(\zeta+1)\tau) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(k+\xi\tau) & \boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(k+(\xi+1)\tau) & \cdots & \boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(k+(\xi+\zeta)\tau) \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{O}_p \\ \boldsymbol{O}_p A^{\tau} \\ \vdots \\ \boldsymbol{O}_p A^{\xi\tau} \end{bmatrix} A^k \begin{bmatrix} \boldsymbol{R}_{\gamma} & A^{\tau}\boldsymbol{R}_{\gamma} & \cdots & A^{\zeta\tau}\boldsymbol{R}_{\gamma} \end{bmatrix} \\ &= \boldsymbol{O}_{\xi} A^k \boldsymbol{R}_{\zeta}. \end{aligned} \tag{1.79}$$

For the case when $k = 0$,

$$\begin{aligned}\boldsymbol{\mathcal{H}}(0) &= \begin{bmatrix} \boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(0) & \boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(\tau) & \cdots & \boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(\zeta\tau) \\ \boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(\tau) & \boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(2\tau) & \cdots & \boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}((\zeta+1)\tau) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}(\xi\tau) & \boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}((\xi+1)\tau) & \cdots & \boldsymbol{\mathcal{R}}_{\boldsymbol{HH}}((\xi+\zeta)\tau) \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{O}_p \\ \boldsymbol{O}_p A^\tau \\ \vdots \\ \boldsymbol{O}_p A^{\xi\tau} \end{bmatrix} \begin{bmatrix} \boldsymbol{R}_\gamma & A^\tau \boldsymbol{R}_\gamma & \cdots & A^{\zeta\tau}\boldsymbol{R}_\gamma \end{bmatrix} \\ &= \boldsymbol{O}_\xi \boldsymbol{R}_\zeta. \end{aligned} \tag{1.80}$$

τ is an integer chosen to chosen to prevent significant overlap of adjacent correlation blocks. The matrices $\boldsymbol{R}_\zeta$ and $\boldsymbol{O}_\xi$ are called the block correlation controllability and observability matrices.

1.8.2 Hankel Norm Approximation

Similarly to the ERA, the ERA/DC process continues with the factorization of the block correlation matrix $\boldsymbol{\mathcal{H}}(0)$ using singular value decomposition so that

$$\boldsymbol{\mathcal{H}}(0) = \boldsymbol{\mathcal{U}}\boldsymbol{\Sigma}\boldsymbol{\mathcal{V}}^\intercal = \begin{bmatrix} \boldsymbol{\mathcal{U}}_n & \boldsymbol{\mathcal{U}}_0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_n & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Sigma}_0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\mathcal{V}}_n^\intercal \\ \boldsymbol{\mathcal{V}}_0^\intercal \end{bmatrix} = \boldsymbol{\mathcal{U}}_n\boldsymbol{\Sigma}_n\boldsymbol{\mathcal{V}}_n^\intercal + \underbrace{\boldsymbol{\mathcal{U}}_0\boldsymbol{\Sigma}_0\boldsymbol{\mathcal{V}}_0^\intercal}_{\simeq \boldsymbol{0}} \simeq \boldsymbol{\mathcal{U}}_n\boldsymbol{\Sigma}_n\boldsymbol{\mathcal{V}}_n^\intercal, \tag{1.81}$$

and

$$\boldsymbol{\mathcal{H}}(0) = \boldsymbol{\mathcal{U}}_n\boldsymbol{\Sigma}_n\boldsymbol{\mathcal{V}}_n^\intercal = \boldsymbol{O}_\xi\boldsymbol{R}_\zeta \Rightarrow \begin{cases} \boldsymbol{O}_\xi = \boldsymbol{\mathcal{U}}_n\boldsymbol{\Sigma}_n^{1/2} \\ \boldsymbol{R}_\zeta = \boldsymbol{\Sigma}_n^{1/2}\boldsymbol{\mathcal{V}}_n^\intercal \end{cases}. \tag{1.82}$$

Again, this choice makes both $\boldsymbol{O}_\xi$ and $\boldsymbol{R}_\zeta$ balanced.

1.8.3 Minimum Realization

From Eq. (1.80) we have directly

$$\boldsymbol{O}_p = \boldsymbol{E}_\gamma^\intercal\boldsymbol{O}_\xi = \boldsymbol{E}_\gamma^\intercal\boldsymbol{\mathcal{U}}_n\boldsymbol{\Sigma}_n^{1/2}. \tag{1.83}$$

From Eq. (1.26), an expression of $\boldsymbol{R}_q$ can be found

$$\boldsymbol{R}_q = \boldsymbol{O}_p^\dagger\boldsymbol{H}(0) = \left[\boldsymbol{E}_\gamma^\intercal\boldsymbol{\mathcal{U}}_n\boldsymbol{\Sigma}_n^{1/2}\right]^\dagger \boldsymbol{H}(0), \tag{1.84}$$

and a realization is shown to be

$$\widehat{A} = \boldsymbol{O}_\xi^\dagger\boldsymbol{\mathcal{H}}(1)\boldsymbol{R}_\zeta^\dagger = \boldsymbol{\Sigma}_n^{-1/2}\boldsymbol{\mathcal{U}}_n^\intercal\boldsymbol{\mathcal{H}}(1)\boldsymbol{\mathcal{V}}_n\boldsymbol{\Sigma}_n^{-1/2}, \tag{1.85a}$$

$$\widehat{B} = \boldsymbol{R}_q\boldsymbol{E}_r = \left[\boldsymbol{E}_\gamma^\intercal\boldsymbol{\mathcal{U}}_n\boldsymbol{\Sigma}_n^{1/2}\right]^\dagger \boldsymbol{H}(0)\boldsymbol{E}_r, \tag{1.85b}$$

$$\widehat{C} = \boldsymbol{E}_m^\intercal\boldsymbol{O}_p = \boldsymbol{E}_m^\intercal\boldsymbol{E}_\gamma^\intercal\boldsymbol{\mathcal{U}}_n\boldsymbol{\Sigma}_n^{1/2}, \tag{1.85c}$$

$$\widehat{D} = h_0. \tag{1.85d}$$

1.9 Conclusion

In this Chapter, several complementary methods have been developed for linear system identification and have been derived using system realization theory, with most methods working well on simulated and test data. The relations between different techniques have been demonstrated and the choice of methods can be made largely on the basis of the final purpose of the identification, for example, control flexible structures. There are many analytical advantages of a linear approximation and, therefore, the search for a linear domain in the system's operational range (if such exists) is crucial. For example, the OKID/ERA method has been successfully applied to identification of real systems, including a linear model of the space shuttle remote manipulator based on a nonlinear simulation code, and the Hubble Space Telescope.

However, the main difficulty in linear system identification applications stems from the interplay of noise and unmodeled dynamics. Noise, finite length of data, and parameters variation are some of the issues that limit the application of such methods and there are many instances when this limitation is significant enough that it becomes necessary to deal with situations where no model in the model set can adequately describe the real system behavior.

In addition, most systems are only linear to a first approximation. Depending on the excitation level, the output is disturbed by nonlinear distortions so that the linearity assumption no longer holds. This immediately limits the application of the results obtained by the linear system identification framework.

Bibliography

[1] Juang, J. N. and R. S. Pappa (1985) "An eigensystem realization algorithm for modal parameter identification and model reduction," *Journal of Guidance Control and Dynamics*, **8**(5), pp. 620–627.
URL `http://doi.aiaa.org/10.2514/3.20031`

[2] Juang, J.-N. and R. S. Pappa (1986) "Effects of noise on modal parameters identified by the Eigensystem Realization Algorithm," *Journal of Guidance, Control, and Dynamics*, **9**(3), pp. 294 –303.

[3] Alvin, K. F. and K. C. Park (1994) "Second Order Structural Identification Procedure via State-Space Based System Identification," *AIAA Journal*, **32**(2), pp. 397–406.

[4] Juang, J. N. (2009) "Generalized Bilinear System Identification," *The Journal of Astronautical Sciences*, **58**(1-2).

[5] Juang, J.-N. (2005) "Continuous-time Bilinear System Identification," *Nonlinear Dynamics*, **39**(1-2), pp. 79–94.

[6] Juang, J.-N., M. Phan, L. G. Horta, and R. W. Longman (1991) "Identification of observer/Kalman filter Markov parameters - Theory and experiments," *Journal of Guidance Control and Dynamics*, **16**(2), pp. 320–329.
URL `http://doi.aiaa.org/10.2514/3.21006`

[7] Juang, J.-N. and R. Longman (1999) *Optimized System Identification*, *Technical Report NASA/TM-1999-209711*, NASA Langley Research Center, Langley Research Center, Hampton VA.

[8] Juang, J.-N. and R. S. Pappa (1985) "An Eigensystem Realization Algorithm (ERA) for Modal Parameter Identification and Model Reduction," *Journal of Guidance, Control, and Dynamics*, **8**(5), pp. 620–627.

[9] Juang, P. M. H. L. G., J.-N. and R. W. Longman (1993) "Identification of Observer/Kalman Filter Parameters: Theory and Experiments," *Journal of Guidance, Control, and Dynamics*, **16**(2), pp. 320–329.

[10] Juang, C. J. E., J.-N. and J. R. Wright (1988) "An Eigensystem Realization Algorithm Using Data Correlation (ERA/DC) for Modal Parameter Identification," *Control-Theory and Advanced Technology*, **4**(1), pp. 5–14.

[11] Juang, J.-N., J. E. Cooper, and J. R. Wright (1988) "An eigensystem realization algorithm using data correlations (ERA/DC) for modal parameter identifcation," *Control-Theory and Advanced Technology*, **4**(1), pp. 5–14.

[12] Anderson, B. D. O. and R. E. Skelton (1988) "The Generation of all q-Markov Covers," *IEEE Transactions on Circuits and Systems*, **46**(375-384).

[13] Liu, K., R. E. Skelton, and J. P. Sharkey (1994) "Modeling Hubble Space Telescope Flight Data by Q-Markov Cover Identification," *Journal of Guidance Control and Dynamics*, **17**(2), pp. 250–256.

[14] Shi, G. and R. E. Skelton (1996) "Iterative Identification and Control Using a Weighted Q-Markov COVER with Measurement Noise," *Signal Processing*, **52**(2), pp. 217–234.

[15] Ho, B. L. and R. E. Kalman (1965) "Effective Construction of Linear State-Variable Models from Input/Output Data," *Proceedings of the 3rd Annual Allerton Conference on Circuit and System Theory*, pp. 449–459.

[16] Chen, L. G., C. W. and J.-N. Juang (1992) "Several Recursive Techniques for Observer/Kalman Filter System Identification From Data," *AIAA Guidance, Navigation and Control Conference*, (Paper No. 92-4386).

[17] Phan, H. L. G. J. J.-N., M. and R. W. Longman (June 1992) "Linear system Identification Via an Asymptotically Stable Observer," *NASA Technical Paper 3164*.