Identification of Linear, Discrete-Time Filters via Realization

Daniel N. Miller and Raymond A. de Callafon

Additional information is available at the end of the chapter

1. Introduction

The realization of a discrete-time, linear, time-invariant (LTI) filter from its impulse response provides insight into the role of linear algebra in the analysis of both dynamical systems and rational functions. For an LTI filter, a sequence of output data measured over some finite period of time may be expressed as the linear combination of the past input and the input measured over that same period. For a finite-dimensional LTI filter, the mapping from past input to future output is a finite-rank linear operator, and the effect of past input, that is, the memory of the system, may be represented as a finite-dimensional vector. This vector is the *state* of the system.

The central idea of realization theory is to first identify the mapping from past input to future output and to then factor it into two parts: a map from the input to the state and another from the state to the output. This factorization guarantees that the resulting system representation is both casual and finite-dimensional; thus it can be physically constructed, or *realized*.

System identification is the science of constructing dynamic models from experimentally measured data. Realization-based identification methods construct models by estimating the mapping from past input to future output based on this measured data. The non-deterministic nature of the estimation process causes this mapping to have an arbitrarily large rank, and so a rank-reduction step is required to factor the mapping into a suitable state-space model. Both these steps must be carefully considered to guarantee unbiased estimates of dynamic systems.

The foundations of realization theory are primarily due to Kalman and first appear in the landmark paper of [1], though the problem is not defined explicitly until [2], which also coins the term "realization" as being the a state-space model of a linear system constructed from an experimentally measured impulse response. It was [3] that introduced the structured-matrix approach now synonymous with the term "realization theory" by re-interpreting a theorem originally due to [4] in a state-space LTI system framework.

Although Kalman's original definition of "realization" implied an identification problem, it was not until [5] proposed rank-reduction by means of the singular-value decomposition

that Ho's method became feasible for use with non-deterministic data sets. The combination of Ho's method and the singular-value decomposition was finally generalized to use with experimentally measured data by Kung in [6].

With the arrival of Kung's method came the birth of what is now known as the field of *subspace identification* methods. These methods use structured matrices of arbitrary input and output data to estimate a state-sequence from the system. The system is then identified from the propagation of the state over time. While many subspace methods exist, the most popular are the Multivariable Output-Error State Space (MOESP) family, due to [7], and the Numerical Algorithms for Subspace State-Space System Identification (N4SID) family, due to [8]. Related to subspace methods is the Eigensystem Realization Algorithm [9], which applies Kung's algorithm to impulse-response estimates, which are typically estimated through an Observer/Kalman Filter Identification (OKID) algorithm [10].

This chapter presents the central theory behind realization-based system identification in a chronological context, beginning with Kronecker's theorem, proceeding through the work of Kalman and Kung, and presenting a generalization of the procedure to arbitrary sets of data. This journey provides an interesting perspective on the original role of linear algebra in the analysis of rational functions and highlights the similarities of the different representations of LTI filters. Realization theory is a diverse field that connects many tools of linear algebra, including structured matrices, the QR-decomposition, the singular-value decomposition, and linear least-squares problems.

2. Transfer-function representations

We begin by reviewing some properties of discrete-time linear filters, focusing on the role of infinite series expansions in analyzing the properties of rational functions. The reconstruction of a transfer function from an infinite impulse response is equivalent to the reconstruction of a rational function from its Laurent series expansion. The reconstruction problem is introduced and solved by forming structured matrices of impulse-response coefficients.

2.1. Difference equations and transfer functions

Discrete-time linear filters are most frequently encountered in the form of difference equations that relate an input signal u_k to an output signal y_k. A simple example is an output y_k determined by a weighted sum of the inputs from u_k to u_{k-m},

$$y_k = b_m u_k + b_{m-1} u_{k-1} + \cdots + b_0 u_{k-m}. \tag{1}$$

More commonly, the output y_k also contains a weighted sum of previous outputs, such as a weighted sum of samples from y_{k-1} to y_{k-n},

$$y_k = b_m u_k + b_{m-1} u_{k-1} + \cdots + b_0 u_{k-m} - a_{n-1} y_{k-1} - a_{n-2} y_{k-2} - \cdots + a_0 y_{k-n}. \tag{2}$$

The impulse response of a filter is the output sequence $g_k = y_k$ generated from an input

$$u_k = \begin{cases} 1 & k = 0, \\ 0 & k \neq 0. \end{cases} \tag{3}$$

The parameters g_k are the impulse-response coefficients, and they completely describe the behavior of an LTI filter through the convolution

$$y_k = \sum_{j=0}^{\infty} g_j u_{k-j}. \tag{4}$$

Filters of type (1) are called finite-impulse response (FIR) filters because g_k is a finite-length sequence that settles to 0 once $k > m$. Filters of type (2) are called infinite impulse response (IIR) filters since generally the impulse response will never completely settle to 0.

A system is stable if a bounded u_k results in a bounded y_k. Because the output of LTI filters is a linear combination of the input and previous output, any input-output sequence can be formed from a linear superposition of other input-output sequences. Hence proving that the system has a bounded output for a single input sequence is necessary and sufficient to prove the stability of an LTI filter. The simplest input to consider is an impulse, and so a suitable definition of system stability is that the absolute sum of the impulse response is bounded,

$$\sum_{k=0}^{\infty} |g_k| < \infty. \tag{5}$$

Though the impulse response completely describes the behavior of an LTI filter, it does so with an infinite number of parameters. For this reason, discrete-time LTI filters are often written as transfer functions of a complex variable z. This enables analysis of filter stability and computation of the filter's frequency response in a finite number of calculations, and it simplifies convolution operations into basic polynomial algebra.

The transfer function is found by grouping output and input terms together and taking the Z-transform of both signals. Let $Y(z) = \sum_{k=-\infty}^{\infty} y_k z^{-k}$ be the Z-transform of y_k and $U(z)$ be the Z-transform of u_k. From the property

$$\mathcal{Z}\left[y_{k-1}\right] = Y(z)z^{-1}$$

the relationship between $Y(z)$ and $U(z)$ may be expressed in polynomials of z as

$$a(z)Y(z) = b(z)U(z).$$

The ratio of these two polynomials is the filter's transfer function

$$G(z) = \frac{b(z)}{a(z)} = \frac{b_m z^m + b_{m-1}z^{m-1} + \cdots + b_1 z + b_0}{z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0}. \tag{6}$$

When $n \geq m$, $G(z)$ is *proper*. If the transfer function is not proper, then the difference equations will have y_k dependent on future input samples such as u_{k+1}. Proper transfer functions are required for causality, and thus all physical systems have proper transfer function representations. When $n > m$, the system is *strictly proper*. Filters with strictly proper transfer functions have no feed-through terms; the output y_k does not depend on u_k, only the preceding input u_{k-1}, u_{k-2}, $\ldots$. In this chapter, we assume all systems are causal and all transfer functions proper.

If $a(z)$ and $b(z)$ have no common roots, then the rational function $G(z)$ is *coprime*, and the order n of $G(z)$ cannot be reduced. Fractional representations are not limited to

single-input-single-output systems. For vector-valued input signals $u_k \in \mathbb{R}^{n_u}$ and output signals $y_k \in \mathbb{R}^{n_y}$, an LTI filter may be represented as an $n_y \times n_u$ matrix of rational functions $G_{ij}(z)$, and the system will have matrix-valued impulse-response coefficients. For simplicity, we will assume that transfer function representations are single-input-single-output, though all results presented here generalize to the multi-input-multi-output case.

2.2. Stability of transfer function representations

Because the effect of $b(z)$ is equivalent to a finite-impulse response filter, the only requirement for $b(z)$ to produce a stable system is that its coefficients be bounded, which we may safely assume is always the case. Thus the stability of a transfer function $G(z)$ is determined entirely by $a(z)$, or more precisely, the roots of $a(z)$. To see this, suppose $a(z)$ is factored into its roots, which are the poles p_i of $G(z)$,

$$G(z) = \frac{b(z)}{\prod_{i=1}^{n}(z - p_i)}. \tag{7}$$

To guarantee a bounded y_k, it is sufficient to study a single pole, which we will denote simply as p. Thus we wish to determine necessary and sufficient conditions for stability of the system

$$G'(z) = \frac{1}{z - p}. \tag{8}$$

Note that p may be complex. Assume that $|z| > |p|$. $G'(z)$ then has the Laurent-series expansion

$$G'(z) = z^{-1} \left(\frac{1}{1 - pz^{-1}} \right) = z^{-1} \sum_{k=0}^{\infty} p^k z^{-k} = \sum_{k=1}^{\infty} p^{k-1} z^{-k}. \tag{9}$$

From the time-shift property of the z-transform, it is immediately clear that the sequence

$$g_k' = \begin{cases} 0 & k = 1, \\ p^{k-1} & k > 1, \end{cases} \tag{10}$$

is the impulse response of $G'(z)$. If we require that (9) is absolutely summable and let $|z| = 1$, the result is the original stability requirement (5), which may be written in terms of p as

$$\sum_{k=1}^{\infty} \left| p^{k-1} \right| < \infty.$$

This is true if and only if $|p| < 1$, and thus $G'(z)$ is stable if and only if $|p| < 1$. Finally, from (7) we may deduce that a system is stable if and only if all the poles of $G(z)$ satisfy the property $|p_i| < 1$.

2.3. Construction of transfer functions from impulse responses

Transfer functions are a convenient way of representing complex system dynamics in a finite number of parameters, but the coefficients of $a(z)$ and $b(z)$ cannot be measured directly. The impulse response of a system can be found experimentally by either direct measurement or from other means such as taking the inverse Fourier transform of a measured frequency response [11]. It cannot, however, be represented in a finite number of parameters. Thus the conversion between transfer functions and impulse responses is an extremely useful tool.

For a single-pole system such as (8), the expansion (9) provides an obvious means of reconstructing a transfer function from a measured impulse response: given any 2 sequential impulse-response coefficients g_k and g_{k+1}, the pole of $G'(z)$ may be found from

$$p = g_k^{-1} g_{k+1}. \tag{11}$$

Notice that this is true for any k, and the impulse response can be said to have a *shift-invariant* property in this respect.

Less clear is the case when an impulse response is generated by a system with higher-order $a(z)$ and $b(z)$. In fact, there is no guarantee that an arbitrary impulse response is the result of a linear system of difference equations at all. For an LTI filter, however, the coefficients of the impulse response exhibit a linear dependence which may be used to not only verify the linearity of the system, but to construct a transfer function representation as well. The exact nature of this linear dependence may be found by forming a structured matrix of impulse response coefficients and examining its behavior when the indices of the coefficients are shifted forward by a single increment, similar to the single-pole case in (11). The result is stated in the following theorem, originally due to Kronecker [4] and adopted from the English translation of [12].

Theorem 1 (Kronecker's Theorem). *Suppose $G(z) : \mathbb{C} \to \mathbb{C}$ is an infinite series of descending powers of z, starting with z^{-1},*

$$G(z) = g_1 z^{-1} + g_2 z^{-2} + g_3 z^{-3} + \cdots = \sum_{k=1}^{\infty} g_k z^{-k}. \tag{12}$$

Assume $G(z)$ is analytic (the series converges) for all $|z| > 1$. Let H be an infinitely large matrix of the form

$$H = \begin{bmatrix} g_1 & g_2 & g_3 & \cdots \\ g_2 & g_3 & g_4 & \cdots \\ g_3 & g_4 & g_5 & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix} \tag{13}$$

Then H has finite rank n if and only if $G(z)$ is a strictly proper, coprime, rational function of degree n with poles inside the unit circle. That is, $G(z)$ has an alternative representation

$$G(z) = \frac{b(z)}{a(z)} = \frac{b_m z^m + b_{m-1} z^{m-1} + \cdots + b_1 z + b_0}{z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0}, \tag{14}$$

in which $m < n$, all roots of $a(z)$ satisfy $|z| < 1$, $a(z)$ and $b(z)$ have no common roots, and we have assumed without loss of generality that $a(z)$ is monic.

To prove Theorem 1, we first prove that for $k > n$, g_k must be linearly dependent on the previous n terms of the series for H to have finite rank.

Theorem 2. *The infinitely large matrix H is of finite rank n if and only if there exists a finite sequence $\alpha_1, \alpha_2, \cdots, \alpha_n$ such that for $k \geq n$,*

$$g_{k+1} = \sum_{j=1}^{n} \alpha_j g_{k-j+1}, \tag{15}$$

and n is the smallest number with this property.

Proof. Let h_k be the row of H beginning with g_k. If H has rank n, then the first $n + 1$ rows of H are linearly dependent. This implies that for some $1 \leq p \leq n$, h_{p+1} is a linear combination of $h_1, \ldots, h_p$, and thus there exists some sequence α_k such that

$$h_{p+1} = \sum_{j=1}^{p} \alpha_j h_{p-j+1}. \tag{16}$$

The structure and infinite size of H imply that such a relationship must hold for all following rows of H, so that for $q \geq 0$

$$h_{q+p+1} = \sum_{j=1}^{p} \alpha_j h_{q+p-j+1}.$$

Hence any row h_k, $k > p$, can be expressed as a linear combination of the previous p rows. Since H has at least n linearly independent rows, $p = n$, and since this applies element-wise, $\text{rank}(H) = n$ implies (15).

Alternatively, (15) implies a relationship of the form (16) exists, and hence $\text{rank}(H) = p$. Since n is the smallest possible p, this implies $\text{rank}(H) = n$. □

We now prove Theorem 1.

Proof. Suppose $G(z)$ is a coprime rational function of the form (14) with series expansion (12), which we know exists, since $G(z)$ is analytic for $|z| < 1$. Without loss of generality, let $m = n - 1$, since we may always let $b_k = 0$ for some k. Hence

$$\frac{b_{n-1}z^{n-1} + b_{n-2}z^{n-2} + \cdots + b_1 z + b_0}{z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0} = g_1 z^{-1} + g_2 z^{-2} + g_3 z^{-3} + \cdots$$

Multiplying both sides by the denominator of the left,

$$b_{n-1}z^{n-1} + b_{n-2}z^{n-2} + \cdots + b_1 z + b_0$$
$$= g_1 z^{n-1} + (g_2 + g_1 a_{n-1})z^{n-2} + (g_3 + g_2 a_{n-1} + g_1 a_{n-2})z^{n-3} + \cdots,$$

and equating powers of z, we find

$$\begin{aligned}
b_{n-1} &= g_1 \\
b_{n-2} &= g_2 + g_1 a_{n-1} \\
b_{n-3} &= g_3 + g_2 a_{n-1} + g_1 a_{n-2} \\
&\vdots \\
b_1 &= g_{n-1} + g_{n-2}a_{n-1} + \cdots + g_1 a_2 \\
b_0 &= g_n + g_{n-1}a_{n-1} + \cdots + g_1 a_1 \\
0 &= g_{k+1} + g_k a_{n-1} + \cdots + g_{k-n+1}a_0 \qquad k \geq n.
\end{aligned} \tag{17}$$

From this, we have, for $k \geq n$,

$$g_{k+1} = \sum_{j=1}^{n} -a_j g_{k-j+1},$$

which not only shows that (15) holds, but also shows that $\alpha_j = -a_j$. Hence by Theorem 2, H must have finite rank.

Conversely, suppose H has finite rank. Then (15) holds, and we may construct $a(z)$ from α_k and $b(z)$ from (17) to create a rational function. This function must be coprime since its order n is the smallest possible. $\square$

The construction in Theorem 1 is simple to extend to the case in which $G(z)$ is only proper and not strictly proper; the additional coefficient b_n is simply the feed-through term in the impulse response, that is, g_0.

A result of Theorem 2 is that given finite-dimensional, full-rank matrices

$$H_k = \begin{bmatrix} g_k & g_{k+1} & \cdots & g_{k+n-1} \\ g_{k+1} & g_{k+2} & \cdots & g_{k+n} \\ \vdots & \vdots & & \vdots \\ g_{k+n-1} & g_{k+n} & \cdots & g_{k+2n-2} \end{bmatrix} \tag{18}$$

and

$$H_{k+1} = \begin{bmatrix} g_{k+1} & g_{k+2} & \cdots & g_{k+n} \\ g_{k+2} & g_{k+3} & \cdots & g_{k+n+1} \\ \vdots & \vdots & & \vdots \\ g_{k+n} & g_{k+n+1} & \cdots & g_{k+2n-1} \end{bmatrix}, \tag{19}$$

the coefficients of $a(z)$ may be calculated as

$$\begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{bmatrix} = H_k^{-1} H_{k+1}. \tag{20}$$

Notice that (11) is in fact a special case of (20). Thus we need only know the first $2n + 1$ impulse-response coefficients to reconstruct the transfer function $G(z)$: $2n$ to form the matrices H_k and H_{k+1} from (18) and (19), respectively, and possibly the initial coefficient g_0 in case of an nth-order $b(z)$.

Matrices with the structure of H are useful enough to have a special name. A *Hankel matrix* H is a matrix constructed from a sequence $\{h_k\}$ so that each element $H_{(j,k)} = h_{j+k}$. For the Hankel matrix in (13), $h_k = g_{k-1}$. H_k also has an interesting property implied by (20): its row space is invariant under shifting of the index k. Because its symmetric, this is also true for its column space. Thus this matrix is also often referred to as being *shift-invariant*.

While (20) provides a potential method of identifying a system from a measured impulse response, this is not a reliable method to use with measured impulse response coefficients that are corrupted by noise. The exact linear dependence of the coefficients will not be identical

for all k, and the structure of (20) will not be preserved. Inverting H_k will also invert any noise on g_k, potentially amplifying high-frequency noise content. Finally, the system order n is required to be known beforehand, which is usually not the case if only an impulse response is available. Fortunately, these difficulties may all be overcome by reinterpreting the results Kronecker's theorem in a state-space framework. First, however, we more carefully examine the role of the Hankel matrix in the behavior of LTI filters.

2.4. Hankel and Toeplitz operators

The Hankel matrix of impulse response coefficients (13) is more than a tool for computing the transfer function representation of a system from its impulse response. It also defines the mapping of past input signals to future output signals. To define exactly what this means, we write the convolution of (4) around sample $k = 0$ in matrix form as

$$
\begin{bmatrix}
\vdots \\
y_{-3} \\
y_{-2} \\
y_{-1} \\
\hline
y_0 \\
y_1 \\
y_2 \\
\vdots
\end{bmatrix}
=
\left[
\begin{array}{cccc|cccc}
\ddots & & & & & & \cdots & 0 \\
\cdots & g_0 & & & & & & \vdots \\
\cdots & g_1 & g_0 & & & & & \\
\cdots & g_2 & g_1 & g_0 & & & & \\
\hline
\cdots & g_3 & g_2 & g_1 & g_0 & & & \\
\cdots & g_4 & g_3 & g_2 & g_1 & g_0 & & \\
\cdots & g_5 & g_4 & g_3 & g_2 & g_1 & g_0 & \\
& \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}
\right]
\begin{bmatrix}
\vdots \\
u_{-3} \\
u_{-2} \\
u_{-1} \\
\hline
u_0 \\
u_1 \\
u_2 \\
\vdots
\end{bmatrix},
$$

where the vectors and matrix have been partitioned into sections for $k < 0$ and $k \geq 0$. The output for $k \geq 0$ may then be split into two parts:

$$
\underbrace{\begin{bmatrix}
y_0 \\
y_1 \\
y_2 \\
\vdots
\end{bmatrix}}_{y_f}
=
\underbrace{\begin{bmatrix}
g_1 & g_2 & g_3 & \cdots \\
g_2 & g_3 & g_4 & \cdots \\
g_3 & g_4 & g_5 & \cdots \\
\vdots & \vdots & \vdots &
\end{bmatrix}}_{H}
\underbrace{\begin{bmatrix}
u_{-1} \\
u_{-2} \\
u_{-3} \\
\vdots
\end{bmatrix}}_{u_p}
+
\underbrace{\begin{bmatrix}
g_0 & & \cdots & 0 \\
g_1 & g_0 & & \vdots \\
g_2 & g_1 & g_0 & \\
\vdots & \vdots & \vdots & \ddots
\end{bmatrix}}_{T}
\underbrace{\begin{bmatrix}
u_0 \\
u_1 \\
u_2 \\
\vdots
\end{bmatrix}}_{u_f},
\qquad (21)
$$

where the subscripts p and f denote "past" and "future," respectively. The system Hankel matrix H has returned to describe the effects of the past input u_p on the future output y_f. Also present is the matrix T, which represents the convolution of future input u_f with the impulse response. Matrices such as T with constant diagonals are called *Toeplitz* matrices.

From (21), it can be seen that H defines the effects of past input on future output. One interpretation of this is that H represents the "memory" of the system. Because H is a linear mapping from u_p to y_f, the induced matrix 2-norm of H, $||H||_2$, can be considered a function norm, and in a sense, $||H||_2$ is a measure of the "gain" of the system. $||H||_2$ is often called the *Hankel-norm* of a system, and it plays an important role in model reduction and in the analysis

of anti-causal systems. More information on this aspect of linear systems can be found in the literature of robust control, for instance, [13].

3. State-space representations

Although transfer functions define system behavior completely with a finite number of parameters and simplify frequency-response calculations, they are cumbersome to manipulate when the input or output is multi-dimensional or when initial conditions must be considered. The other common representation of LTI filters is the state-space form

$$
\begin{aligned}
x_{k+1} &= Ax_k + Bu_k \\
y_k &= Cx_k + Du_k,
\end{aligned}
\tag{22}
$$

in which $x_k \in \mathbb{R}^n$ is the system state. The matrices $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times n_u}$, $C \in \mathbb{R}^{n_y \times n}$, and $D \in \mathbb{R}^{n_y \times n_u}$ completely parameterize the system. Only D uniquely defines the input-output behavior; any nonsingular matrix T' may be used to change the state basis via the relationships

$$
x' = T'x \qquad A' = T'AT'^{-1} \qquad B' = T'B \qquad C' = CT'^{-1}.
$$

The Z-transform may also be applied to the state-space equations (22) to find

$$
\begin{aligned}
\mathcal{Z}[x_{k+1}] &= A\mathcal{Z}[x_k] + B\mathcal{Z}[u_k] &\qquad \Rightarrow \qquad & X(z)z = AX(z) + BU(z) \\
\mathcal{Z}[y_k] &= C\mathcal{Z}[x_k] + D\mathcal{Z}[u_k] &\qquad \Rightarrow \qquad & Y(z) = CX(z) + DU(z)
\end{aligned}
$$

$$
\frac{Y(z)}{U(z)} = G(z) \qquad G(z) = C\,(zI - A)^{-1}\,B + D,
\tag{23}
$$

and thus, if (22) is the state-space representation of the single-variable system (6), then $a(z)$ is the characteristic polynomial of A, $\det(zI - A)$.

Besides clarifying the effect of initial conditions on the output, state-space representations are inherently causal, and (23) will always result in a proper system (strictly proper if $D = 0$). For this reason, state-space representations are often called *realizable* descriptions; while the forward-time-shift of z is an inherently non-causal operation, state-space systems may always be constructed in reality.

3.1. Stability, controllability, and observability of state-space representations

The system impulse response is simple to formulate in terms of the state-space parameters by calculation of the output to a unit impulse with $x_0 = 0$:

$$
g_k = \begin{cases} D & k = 0, \\ CA^{k-1}B & k > 0 \end{cases}.
\tag{24}
$$

Notice the similarity of (10) and (24). In fact, from the eigenvalue decomposition of A,

$$
A = V\Lambda V^{-1},
$$

we find

$$
\sum_{k=1}^{\infty} |g_k| = \sum_{k=1}^{\infty} \left| CA^{k-1}B \right| = \sum_{k=1}^{\infty} |CV|\,(|\Lambda^{k-1}|)\,\left| V^{-1}B \right|.
$$

The term $|\Lambda^{k-1}|$ will only converge if the largest eigenvalue of A is within the unit circle, and thus the condition that all eigenvalues λ_i of A satisfy $|\lambda_i| < 1$ is a necessary and sufficient condition for stability.

For state-space representations, there is the possibility that a combination of A and B will result in a system for which x_k cannot be entirely controlled by the input u_k. Expressing x_k in a matrix-form similar to (21) as

$$x_k = \mathcal{C} \begin{bmatrix} u_{k-1} \\ u_{k-2} \\ u_{k-3} \\ \vdots \end{bmatrix}, \qquad \mathcal{C} = \begin{bmatrix} B & AB & A^2B & \cdots \end{bmatrix} \tag{25}$$

demonstrates that x_k is in subspace $\mathbb{R}^n$ if and only if $\mathcal{C}$ has rank n. $\mathcal{C}$ is the *controllability matrix* and the system is *controllable* if it has full row rank.

Similarly, the state x_k may not uniquely determine the output for some combinations of A and C. Expressing the evolution of the output as a function of the state in matrix-form as

$$\begin{bmatrix} y_k \\ y_{k+1} \\ y_{k+2} \\ \vdots \end{bmatrix} = \mathcal{O} x_k, \qquad \mathcal{O} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix}$$

demonstrates that there is no nontrivial null space in the mapping from x_k to y_k if and only if $\mathcal{O}$ has rank n. $\mathcal{O}$ is the *observability matrix* and the system is *observable* if it has full column rank.

Systems that are both controllable and observable are called *minimal*, and for minimal systems, the dimension n of the state variable cannot be reduced. In the next section we show that minimal state-space system representations convert to coprime transfer functions that are found through (23).

3.2. Construction of state-space representations from impulse responses

The fact that the denominator of $G(z)$ is the characteristic polynomial of A not only allows for the calculation of a transfer function from a state-space representation, but provides an alternative version of Kronecker's theorem for state-space systems, known as the Ho-Kalman Algorithm [3]. From the Caley-Hamilton theorem, if $a(z)$ is the characteristic polynomial of A, then $a(A) = 0$, and

$$CA^k a(A)B = CA^k \left(A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 \right) B$$

$$= CA^{k+n}B + \sum_{j=0}^{n-1} a_j CA^{k+j}B,$$

which implies

$$CA^{k+n}B = -\sum_{j=0}^{n-1} a_j CA^{k+j}B. \tag{26}$$

Indeed, substitution of (24) into (26) and rearrangement of the indices leads to (15). Additionally, substitution of (24) into the product of $\mathcal{O}$ and $\mathcal{C}$ shows that

$$\mathcal{O}\mathcal{C} = \begin{bmatrix} CB & CAB & CA^2B & \cdots \\ CAB & CA^2B & CA^3B & \cdots \\ CA^2B & CA^3B & CA^4B & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix} = \begin{bmatrix} g_1 & g_2 & g_3 & \cdots \\ g_2 & g_3 & g_4 & \cdots \\ g_3 & g_4 & g_5 & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix} = H,$$

which confirms our previous statement that H effectively represents the memory of the system. Because

$$\text{rank}(H) = \min\{\text{rank}(\mathcal{O}),\ \text{rank}(\mathcal{C})\},$$

we see that $\text{rank}(H) = n$ implies the state-space system (22) is minimal.

If the entries of H are shifted forward by one index to form

$$\overline{H} = \begin{bmatrix} g_2 & g_3 & g_4 & \cdots \\ g_3 & g_4 & g_5 & \cdots \\ g_4 & g_5 & g_6 & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix},$$

then once again substituting (24) reveals

$$\overline{H} = \mathcal{O}A\mathcal{C}. \tag{27}$$

Thus the row space and column space of H are invariant under a forward-shift of the indices, implying the same shift-invariant structure seen in (20).

The appearance of A in (27) hints at a method for constructing a state-space realization from an impulse response. Suppose the impulse response is known exactly, and let H_r be a finite slice of H with r block rows and L columns,

$$H_r = \begin{bmatrix} g_1 & g_2 & g_3 & \cdots & g_L \\ g_2 & g_3 & g_4 & \cdots & g_{L+1} \\ g_3 & g_4 & g_5 & \cdots & g_{L+2} \\ \vdots & \vdots & \vdots & & \vdots \\ g_{r-1} & g_r & g_{r+1} & \cdots & g_{r+L-1} \end{bmatrix}.$$

Then any appropriately dimensioned factorization

$$H_r = \mathcal{O}_r \mathcal{C}_L = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{r-1} \end{bmatrix} \begin{bmatrix} B & AB & A^2B & \cdots & A^{L-1}B \end{bmatrix} \tag{28}$$

may be used to find A for some arbitrary state basis as

$$A = (\mathcal{O}_r)^\dagger \overline{H}_r (\mathcal{C}_L)^\dagger \tag{29}$$

where $\overline{H}_r$ is H_r with the indices shifted forward once and $(\cdot)^\dagger$ is the Moore-Penrose pseudoinverse. C taken from the first block row of $\mathcal{O}_r$, B taken from the first block column of $\mathcal{C}_L$, and D taken from g_0 then provides a complete and minimal state-space realization from an impulse response. Because H_r has rank n and $\det(zI - A)$ has degree n, we know from Kronecker's theorem that $G(z)$ taken from (23) will be coprime.

However, as mentioned before, the impulse response of the system is rarely known exactly. In this case only an estimate $\hat{H}_r$ with a non-deterministic error term is available:

$$\hat{H}_r = H_r + E.$$

Because E is non-deterministic, $\hat{H}$ will always have full rank, regardless of the number of rows r. Thus n cannot be determined from examining the rank of $\hat{H}$, and even if n is known beforehand, a factorization (28) for $r > n$ will not exist. Thus we must find a way of reducing the rank of $\hat{H}_r$ in order to find a state-space realization.

3.3. Rank-reduction of the Hankel matrix estimate

If $\hat{H}_r$ has full rank, or if n is unknown, its rank must be reduced prior to factorization. The obvious tool for reducing the rank of matrices is the *singular-value decomposition* (SVD). Assume for now that n is known. The SVD of $\hat{H}_r$ is

$$\hat{H}_r = U\Sigma V^T$$

where U and V^T are orthogonal matrices and Σ is a diagonal matrix containing the nonnegative *singular values* σ_i ordered from largest to smallest. The SVD for a matrix is unique and guaranteed to exist, and the number of nonzero singular values of a matrix is equal to its rank [14].

Because U and V^T are orthogonal, the SVD satisfies

$$\hat{H}_r = \left|\left| U\Sigma V^T \right|\right|_2 = ||\Sigma||_2 = \sigma_1 \tag{30}$$

where $||\cdot||_2$ is the induced matrix 2-norm, and

$$\hat{H}_r = \left|\left| U\Sigma V^T \right|\right|_F = ||\Sigma||_F = \left(\sum_i^l \sigma_i^2 \right)^{1/2} \tag{31}$$

where $||\cdot||_F$ is the Frobenius norm. Equation (30) also shows that the Hankel norm of a system is the maximum singular value of H_r. From (30) and (31), we can directly see that if the SVD of H_r is partitioned into

$$\hat{H}_r = \begin{bmatrix} U_n & U_s \end{bmatrix} \begin{bmatrix} \Sigma_n & 0 \\ 0 & \Sigma_s \end{bmatrix} \begin{bmatrix} V_n^T \\ V_s^T \end{bmatrix},$$

where U_n is the first n columns of U, Σ_n is the upper-left $n \times n$ block of Σ, and V_n^T is the first n rows of V^T, the solution to the rank-reduction problem is [14]

$$Q = \underset{\mathrm{rank}(Q)=n}{\arg\min} \left|\left| Q - \hat{H}_r \right|\right|_2 = \underset{\mathrm{rank}(Q)=n}{\arg\min} \left|\left| Q - \hat{H}_r \right|\right|_F = U_n \Sigma_n V_n^T.$$

Additionally, the error resulting from the rank reduction is

$$e = \left|\left|Q - \hat{H}_r\right|\right|_2 = \sigma_{n+1},$$

which suggests that if the rank of H_r is not known beforehand, it can be determined by examining the nonzero singular values in the deterministic case or by searching for a significant drop-off in singular values if only a noise-corrupted estimate is available.

3.4. Identifying the state-space realization

From a rank-reduced $\hat{H}_r$, any factorization

$$\hat{H}_r = \hat{\mathcal{O}}_r \hat{\mathcal{C}}_L$$

can be used to estimate $\mathcal{O}_r$ and $\mathcal{C}_L$. The error in the state-space realization, however, will depend on the chosen state basis. Generally we would like to have a state variable with a norm $||x_k||_2$ in between $||u_k||_2$ and $||y_k||_2$. As first proposed in [5], choosing the factorization

$$\hat{\mathcal{O}}_r = U_n \Sigma_n^{1/2} \qquad \text{and} \qquad \hat{\mathcal{C}}_L = \Sigma_n^{1/2} V_n^T \tag{32}$$

results in

$$\left|\left|\hat{\mathcal{O}}_r\right|\right|_2 = \left|\left|\hat{\mathcal{C}}_L\right|\right|_2 = \sqrt{\left|\left|\hat{H}_r\right|\right|_2}, \tag{33}$$

and thus, from a functional perspective, the mappings from input to state and state to output will have equal magnitudes, and each entry of the state vector x_k will have similar magnitudes. State-space realizations that satisfy (33) are sometimes called *internally balanced* realizations [11]. (Alternative definitions of a "balanced" realization exist, however, and it is generally wise to verify the definition in each context.)

Choosing the factorization (32) also simplifies computation of the estimate $\hat{A}$, since

$$\begin{aligned} \hat{A} &= \left(\hat{\mathcal{O}}_r\right)^\dagger \hat{\bar{H}}_r \left(\hat{\mathcal{C}}_L\right)^\dagger \\ &= \Sigma_n^{-1/2} U_n^T \hat{\bar{H}}_r V_n \Sigma_n^{-1/2}. \end{aligned}$$

By estimating $\hat{B}$ as the first block column of $\hat{\mathcal{C}}_L$, $\hat{C}$ as the first block row of $\hat{\mathcal{O}}_L$, and $\hat{D}$ as g_0, a complete state-space realization $(\hat{A},\ \hat{B},\ \hat{C},\ \hat{D})$ is identified from this method.

3.5. Pitfalls of direct realization from an impulse response

Even though the rank-reduction process allows for realization from a noise-corrupted estimate of an impulse response, identification methods that generate a system estimate from a Hankel matrix constructed from an estimated impulse response have numerous difficulties when applied to noisy measurements. Measuring an impulse response directly is often infeasible; high-frequency damping may result in a measurement that has a very brief response before the signal-to-noise ratio becomes prohibitively small, and a unit pulse will often excite high-frequency nonlinearities that degrade the quality of the resulting estimate.

Taking the inverse Fourier transform of the frequency response guarantees that the estimates of the Markov parameters will converge as the dataset grows only so long as the input is broadband. Generally input signals decay in magnitude at higher frequencies, and calculation

of the frequency response by inversion of the input will amplify high-frequency noise. We would prefer an identification method that is guaranteed to provide a system estimate that converges to the true system as the amount of data measured increases and that avoids inverting the input. Fortunately, the relationship between input and output data in (21) may be used to formulate just such an identification procedure.

4. Realization from input-output data

To avoid the difficulties in constructing a system realization from an estimated impulse response, we will form a realization-based identification procedure applicable to measured input-output data. To sufficiently account for non-deterministic effects in measured data, we add a noise term $v_k \in \mathbb{R}^{n_y}$ to the output to form the noise-perturbed state-space equations

$$
\begin{aligned}
x_{k+1} &= Ax_k + Bu_k \\
y_k &= Cx_k + Du_k + v_k.
\end{aligned}
\tag{34}
$$

We assume that the noise signal v_k is generated by a stationary stochastic process, which may be either white or colored. This includes the case in which the state is disturbed by process noise, so that the noise process may have the same poles as the deterministic system. (See [15] for a thorough discussion of representations of noise in the identification context.)

4.1. Data-matrix equations

The goal is to construct a state-space realization using the relationships in (21), but doing so requires a complete characterization of the row space of H_r. To this end, we expand a finite-slice of the future output vector to form a block-Hankel matrix of output data with r block rows,

$$
Y = \begin{bmatrix}
y_0 & y_1 & y_2 & \cdots & y_L \\
y_1 & y_2 & y_3 & \cdots & y_{L+1} \\
y_2 & y_3 & y_4 & \cdots & y_{L+2} \\
\vdots & \vdots & \vdots & & \vdots \\
y_{r-1} & y_r & y_{r+1} & \cdots & y_{r+L-1}
\end{bmatrix}.
$$

This matrix is related to a block-Hankel matrix of future input data

$$
U_f = \begin{bmatrix}
u_0 & u_1 & u_2 & \cdots & u_L \\
u_1 & u_2 & u_3 & \cdots & u_{L+1} \\
u_2 & u_3 & u_4 & \cdots & u_{L+2} \\
\vdots & \vdots & \vdots & & \vdots \\
u_{r-1} & u_r & u_{r+1} & \cdots & u_{r+L-1}
\end{bmatrix},
$$

a block-Toeplitz matrix of past input data

$$
U_p = \begin{bmatrix}
u_{-1} & u_0 & u_1 & \cdots & u_{L-1} \\
u_{-2} & u_{-1} & u_0 & \cdots & u_{L-2} \\
u_{-3} & u_{-2} & u_{-1} & \cdots & u_{L-3} \\
\vdots & \vdots & \vdots & & \vdots
\end{bmatrix},
$$

a finite-dimensional block-Toeplitz matrix

$$
T = \begin{bmatrix}
g_0 & & & & \cdots & 0 \\
g_1 & g_0 & & & & \vdots \\
g_2 & g_1 & g_0 & & & \\
\vdots & \vdots & \vdots & \ddots & & \\
g_{r-1} & g_{r-2} & g_{r-3} & \cdots & & g_0
\end{bmatrix},
$$

the system Hankel matrix H, and a block-Hankel matrix V formed from noise data v_k with the same indices as Y by the equation

$$
Y = HU_p + TU_f + V. \tag{35}
$$

If the entries of Y_f are shifted forward by one index to form

$$
\overline{Y} = \begin{bmatrix}
y_1 & y_2 & y_3 & \cdots & y_{L+1} \\
y_2 & y_3 & y_4 & \cdots & y_{L+2} \\
y_3 & y_4 & y_5 & \cdots & y_{L+3} \\
\vdots & \vdots & \vdots & & \vdots \\
y_r & y_{r+1} & y_{r+2} & \cdots & y_{r+L}
\end{bmatrix},
$$

then $\overline{Y}_f$ is related to the shifted system Hankel matrix $\overline{H}$, the past input data U_p, T with a block column appended to the left, and U_f with a block row appended to the bottom,

$$
\overline{T} = \begin{bmatrix}
\begin{matrix} g_1 \\ g_2 \\ g_3 \\ \vdots \\ g_r \end{matrix} & T
\end{bmatrix}, \qquad
\overline{U}_f = \left[\begin{array}{c} U_f \\ \hline u_r \; u_{r+1} \; u_{r+2} \; \cdots \; u_{r+L} \end{array} \right],
$$

and a block-Hankel matrix $\overline{V}$ of noise data v_k with the same indices as $\overline{Y}$ by the equation

$$
\overline{Y} = \overline{H}U_p + \overline{T}\,\overline{U}_f + \overline{V}. \tag{36}
$$

From (25), the state is equal to the column vectors of U_p multiplied by the entries of the controllability matrix $\mathcal{C}$, which we may represent as the block-matrix

$$
X = \begin{bmatrix} x_0 \; x_1 \; x_2 \; \cdots \; x_L \end{bmatrix} = \mathcal{C}U_p,
$$

which is an alternative means of representing the memory of the system at sample 0, 1, ... The two data matrix equations (35) and (36) may then be written as

$$
Y = \mathcal{O}_r X + TU_f + V, \tag{37}
$$
$$
\overline{Y} = \mathcal{O}_r AX + \overline{T}\,\overline{U}_f + \overline{V}. \tag{38}
$$

Equation (37) is basis for the field of system identification methods known as subspace methods. Subspace identification methods typically fall into one of two categories. First,

because a shifted observability matrix

$$\overline{\mathcal{O}} = \begin{bmatrix} CA \\ CA^2 \\ CA^3 \\ \vdots \end{bmatrix},$$

satisfies

$$\mathrm{im}(\mathcal{O}) = \mathrm{im}(\overline{\mathcal{O}}),$$

where $\mathrm{im}(\cdot)$ of denotes the row space (often called the "image"), the row-space of $\mathcal{O}$ is shift-invariant, and A may be identified from estimates $\mathcal{O}_r$ and $\overline{\mathcal{O}}_r$ as

$$\hat{A} = \mathcal{O}_r^\dagger \hat{\overline{\mathcal{O}}}_r.$$

Alternatively, because a forward-propagated sequence of states

$$\overline{X} = AX$$

satisfies

$$\mathrm{im}(X^T) = \mathrm{im}(\overline{X}^T),$$

the column-space of X is shift-invariant, and A may be identified from estimates $\hat{X}$ and $\hat{\overline{X}}$ as

$$\hat{A} = \hat{\overline{X}} \hat{X}^\dagger.$$

In both instances, the system dynamics are estimated by propagating the indices forward by one step and examining a propagation of linear dynamics, not unlike (20) from Kronecker's theorem. Details of these methods may be found in [16] and [17]. In the next section we present a system identification method that constructs system estimates from the shift-invariant structure of Y itself.

4.2. Identification from shift-invariance of output measurements

Equations (37) and (38) still contain the effects of the future input in $\overline{U}_f$. To remove these effects from the output, we must first add some assumptions about $\overline{U}_f$. First, we assume that $\overline{U}_f$ has full row rank. This is true for any U_f with a smooth frequency response or if $\overline{U}_f$ is generated from some pseudo-random sequence. Next, we assume that the initial conditions in X do not somehow cancel out the effects of future input. A sufficient condition for this is to require

$$\mathrm{rank}\left(\begin{bmatrix} X \\ \overline{U}_f \end{bmatrix}\right) = n + rn_u$$

to have full row rank. Although these assumptions might appear restrictive at first, since it is impossible to verify without knowledge of X, it is generally true with the exception of some pathological cases.

Next, we form the null-space projector matrix

$$\Pi = I_{L+1} - \overline{U}_f^T \left(\overline{U}_f \overline{U}_f^T\right)^{-1} \overline{U}_f, \tag{39}$$

which has the property

$$\overline{U}_f \Pi = 0.$$

We know the inverse of $(\overline{U}_f \overline{U}_f^T)$ exists, since we assume $\overline{U}_f$ has full row rank. Projector matrices such as (39) have many interesting properties. Their eigenvalues are all 0 or 1, and if they are symmetric, they separate the subspace of real vectors — in this case, vectors in $\mathbb{R}^{L+1}$ — into a subspace and its orthogonal complement. In fact, it is simple to verify that the null space of $\overline{U}_f$ contains the null space of U_f as a subspace, since

$$\overline{U}_f \Pi = \left[\frac{U_f}{\cdots} \right] \Pi = 0.$$

Thus multiplication of (37) and (38) on the right by Π results in

$$Y\Pi = \mathcal{O}_r X\Pi + V\Pi, \tag{40}$$

$$\overline{Y}\Pi = \mathcal{O}_r A X\Pi + \overline{V}\Pi. \tag{41}$$

It is also unnecessary to compute the projected products $Y\Pi$ and $\overline{Y}\Pi$ directly, since from the QR-decomposition

$$\begin{bmatrix} \overline{U}^T & Y^T \end{bmatrix} = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix},$$

we have

$$Y = R_{12}^T Q_1^T + R_{22}^T Q_2^T \tag{42}$$

and $U = R_{11}^T Q_1^T$. Substitution into (39) reveals

$$\Pi = I - Q_1 Q_1^T. \tag{43}$$

Because the columns of Q_1 and Q_2 are orthogonal, multiplication of (42) on the right by (43) results in

$$Y\Pi = R_{22}^T Q_2^T.$$

A similar result holds for $\overline{Y}\Pi$. Taking the QR-decomposition of the data can alternatively be thought of as using the principle of superposition to construct new sequences of input-output data through a Gram-Schmidt-type orthogonalization process. A detailed discussion of this interpretation can be found in [18].

Thus we have successfully removed the effects of future input on the output while retaining the effects of the past, which is the foundation of the realization process. We still must account for non-deterministic effects in V and $\overline{V}$. To do so, we look for some matrix Z such that

$$V\Pi Z^T \to 0,$$

$$\overline{V}\Pi Z^T \to 0.$$

This requires the content of Z to be statistically independent of the process that generates v_k. The input u_k is just such a signal, so long as the filter output is not a function of the input — that is, the data was measured in open-loop operation,. If we begin measuring input before

$k = 0$ at some sample $k = -\zeta$ and construct Z as a block-Hankel matrix of past input data,

$$Z = \frac{1}{L} \begin{bmatrix} u_{-\zeta} & u_{-\zeta+1} & u_{-\zeta+2} & \cdots & u_{-\zeta+L} \\ u_{-\zeta+1} & u_{-\zeta+2} & u_{-\zeta+3} & \cdots & u_{-\zeta+L+1} \\ u_{-\zeta+2} & u_{-\zeta+3} & u_{-\zeta+4} & \cdots & u_{-\zeta+L+2} \\ \vdots & \vdots & \vdots & & \vdots \\ u_{-1} & u_0 & u_1 & \cdots & u_{r+L-2} \end{bmatrix},$$

then multiplication of (40) and (41) on the right by Z^T results in

$$Y\Pi Z^T \rightarrow \mathcal{O}_r X \Pi Z^T, \tag{44}$$

$$\overline{Y}\Pi Z^T \rightarrow \mathcal{O}_r A \Pi Z^T, \tag{45}$$

as $L \rightarrow \infty$. Note the term $\frac{1}{L}$ in Z is necessary to keep (44) and (45) bounded.

Finally we are able to perform our rank-reduction technique directly on measured data without needing to first estimate the impulse response. From the SVD

$$Y\Pi Z^T = U\Sigma V^T,$$

we may estimate the order n by looking for a sudden decrease in singular values. From the partitioning

$$Y\Pi Z^T = \begin{bmatrix} U_n & U_s \end{bmatrix} \begin{bmatrix} \Sigma_n & 0 \\ 0 & \Sigma_s \end{bmatrix} \begin{bmatrix} V_n^T \\ V_s^T \end{bmatrix},$$

we may estimate $\mathcal{O}_r$ and $X\Pi Z^T$ from the factorization

$$\hat{\mathcal{O}}_r = U_n \Sigma_n^{1/2} \quad \text{and} \quad \hat{X}\Pi Z^T = \Sigma_n^{1/2} V_n^T.$$

A may then be estimated as

$$\hat{A} = \left(\hat{\mathcal{O}}_r\right)^\dagger \overline{Y}\Pi Z^T \left(\hat{X}\Pi Z^T\right)^\dagger = \Sigma_n^{-1/2} U_n^T \overline{Y}\Pi Z^T V_n \Sigma_n^{-1/2}$$

$$\approx \left(\mathcal{O}_r\right)^\dagger \left(\overline{H}U_p\Pi\right)\left(\mathcal{C}_L U_p\Pi\right)^\dagger \approx \left(\mathcal{O}_r\right)^\dagger \overline{H}\left(\mathcal{C}_L\right)^\dagger.$$

And so we have returned to our original relationship (29).

While C may be estimated from the top block row of $\hat{\mathcal{O}}_r$, our projection has lost the column space of H_r that we previously used to estimate B, and initial conditions in X prevent us from estimating D directly. Fortunately, if A and C are known, then the remaining terms B, D, and an initial condition x_0 are linear in the input output data, and may be estimated by solving a linear-least-squares problem.

4.3. Estimation of B, D, and $x0$

The input-to-state terms B and D may be estimated by examining the convolution with the state-space form of the impulse response. Expanding (24) with the input and including an initial condition x_0 results in

$$y_k = CA^k x_0 + \sum_{j=0}^{k-1} CA^{k-j-1} Bu_j + Du_k + v_k. \tag{46}$$

Factoring out B and D on the right provides

$$y_k = CA^k x_0 + \left(\sum_{j=0}^{k-1} u_k^T \otimes CA^{k-j-1} \right) \text{vec}(B) + \left(u_k^T \otimes I_{n_y} \right) \text{vec}(D) + v_k,$$

in which $\text{vec}(\cdot)$ is the operation that stacks the columns of a matrix on one another, $\otimes$ is the (coincidentally named) Kronecker product, and we have made use of the identity

$$\text{vec}(AXB) = (B^T \otimes A)\text{vec}(X).$$

Grouping the unknown terms together results in

$$y_k = \left[CA^k \ \ \sum_{j=0}^{k-1} u_k^T \otimes CA^{k-j-1} \ \ u_k^T \otimes I_{n_y} \right] \begin{bmatrix} x_0 \\ \text{vec}(B) \\ \text{vec}(D) \end{bmatrix} + v_k.$$

Thus by forming the regressor

$$\phi_k = \left[\hat{C}\hat{A}^k \ \ \sum_{j=0}^{k-1} u_k^T \otimes \hat{C}\hat{A}^{k-j-1} \ \ u_k^T \otimes I_{n_y} \right]$$

from the estimates $\hat{A}$ and $\hat{C}$, estimates of B and D may be found from the least-squares solution of the linear system of N equations

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} \phi_0 \\ \phi_1 \\ \phi_2 \\ \vdots \\ \phi_N \end{bmatrix} \begin{bmatrix} \hat{x}_0 \\ \text{vec}(\hat{B}) \\ \text{vec}(\hat{D}) \end{bmatrix}.$$

Note that N is arbitrary and does not need to be related in any way to the indices of the data matrix equations. This can be useful, since for large-dimensional systems, the regressor ϕ_k may become very computationally expensive to compute.

5. Conclusion

Beginning with the construction of a transfer function from an impulse response, we have constructed a method for identification of state-space realizations of linear filters from measured input-output data, introducing the fundamental concepts of realization theory of linear systems along the way.   Computing a state-space realization from measured input-output data requires many tools of linear algebra: projections and the QR-decomposition, rank reduction and the singular-value decomposition, and linear least squares. The principles of realization theory provide insight into the different representations of linear systems, as well as the role of rational functions and series expansions in linear algebra.

Author details

Daniel N. Miller and Raymond A. de Callafon
University of California, San Diego, USA

6. References

[1] Rudolf E. Kalman. On the General Theory of Control Systems. In *Proceedings of the First International Congress of Automatic Control*, Moscow, 1960. IRE.

[2] Rudolf E. Kalman. Mathematical Description of Linear Dynamical Systems. *Journal of the Society for Industrial and Applied Mathematics, Series A: Control*, 1(2):152, July 1963.

[3] B. L. Ho and Rudolf E. Kalman. Effective construction of linear state-variable models from input/output functions. *Regelungstechnik*, 14:545–548, 1966.

[4] Leopold Kronecker. Zur Theorie der Elimination einer Variabeln aus zwei Algebraischen Gleichungen. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 535–600, 1881.

[5] Paul H. Zeiger and Julia A. McEwen. Approximate Linear Realizations of Given Dimension via Ho's Algorithm. *IEEE Transactions on Automatic Control*, 19(2):153 – 153, 1971.

[6] Sun-Yuan Kung. A new identification and model reduction algorithm via singular value decomposition. In *Proceedings of the 12th Asilomar Conference on Circuits, Systems, and Computers*, pages 705–714. IEEE, 1978.

[7] Michel Verhaegen. Identification of the deterministic part of MIMO state space models given in innovations form from input-output data. *Automatica*, 30(1):61–74, January 1993.

[8] Peter Van Overschee and Bart De Moor. N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica*, 30(1):75–93, January 1994.

[9] Jer-Nan Juang and R. S. Pappa. An Eigensystem Realization Algorithm (ERA) for Modal Parameter Identification and Model Reduction. *JPL Proc. of the Workshop on Identification and Control of Flexible Space Structures*, 3:299–318, April 1985.

[10] Jer-Nan Juang, Minh Q. Phan, Lucas G. Horta, and Richard W. Longman. Identification of Observer/Kalman Filter Markov Parameters: Theory and Experiments. *Journal of Guidance, Control, and Dynamics*, 16(2):320–329, 1993.

[11] Chi-Tsong Chen. *Linear System Theory and Design*. Oxford University Press, New York, 1st edition, 1984.

[12] Felix R. Gantmacher. *The Theory of Matrices - Volume Two*. Chelsea Publishing Company, New York, 1960.

[13] Kemin Zhou, John C. Doyle, and Keith Glover. *Robust and Optimal Control*. Prentice Hall, August 1995.

[14] G.H. Golub and C.F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, Maryland, USA, third edition, 1996.

[15] Lennart Ljung. *System Identification: Theory for the User*. PTR Prentice Hall Information and System Sciences. Prentice Hall PTR, Upper Saddle River, NJ, 2nd edition, 1999.

[16] Michel Verhaegen and Vincent Verdult. *Filtering and System Identification: A Least Squares Approach*. Cambridge University Press, New York, 1 edition, May 2007.

[17] Peter Van Overschee and Bart De Moor. *Subspace Identification for Linear Systems: Theory, Implementation, Applications*. Kluwer Academic Publishers, London, 1996.

[18] Tohru Katayama. Role of LQ Decomposition in Subspace Identification Methods. In Alessandro Chiuso, Stefano Pinzoni, and Augusto Ferrante, editors, *Modeling, Estimation and Control*, pages 207–220. Springer Berlin / Heidelberg, 2007.