SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

State-Space Realization and Model-Order Reduction: Modeling under Complexity Constraints

av

Shamiur Rahman Ramim

2026 - No K23

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 106 91 STOCKHOLM

State-Space Realization and Model-Order Reduction: Modeling under Complexity Constraints

Shamiur Rahman Ramim

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Yishao Zhou

2026

Abstract

This thesis deals with the problems of (1) realization (2) partial realization and (3) approximate realization. Focus is directed towards the partial realization problem, which asks for an internal model description for a linear dynamical system provided we only know partial information about its input-output behavior. We study two aspects of this problem: construction and complexity of models, which is measured by the number of equations needed to describe the system. The construction of an internal model description from partially known input-output behavior is fairly trivial but the construction of a simple one is not. We emphasize two methods: one where we treat the problem as one of interpolation and one where we decompose an already known model to a simpler one. If our simplest model is still too complex for any simulation, analysis or controller-synthesis to be made, we turn to the problem of approximate realization in which we find models of lower complexity that approximate the behavior. We treat this approximation problem as a problem of structured low-rank approximation. This all can be seen as a model-order reduction procedure where we make the model smaller while keeping – or in this case, approximating – some desirable attribute.

Sammanfattning

I denna uppsats hanteras problemen om (1) realisation, och (2) partiell realisation, (3) approximativ realisation. Fokuset ligger på det partiella realisationsproblemet som frågar efter en intern modell för ett linjärt dynamiskt system givet ofullständig information om dess in- och utdataförhållande. Vi studerar två aspekter av detta problem: konstruktion och komplexitet av modeller, som mäts i antal ekvationer som krävs för att beskriva systemet. Det är ganska enkelt att konstruera en intern modell från ofullständig in- och utdata, men att konstruera en enkel modell är inte det. Vi betonar två metoder, en där problemet omformuleras till ett om interpolering och en där vi bryter ner en modell till en mindre som ger samma in- och utdataförhållande. Om den enklaste modellen för ett system fortfarande är för komplex så vänder vi oss till modeller av lägre komplexitet som approximerar vårt ofullständiga in- och utdataförhållande. Vi behandlar detta som ett problem om lågrangsapproximation av strukturerade matriser. Allt detta i vår kontext kan ses som en typ av modellreduktion där vi förenklar vår modell medan vi behåller – eller som i detta fall, approximerar – vissa önskvärda egenskaper.

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Yishao Zhou, for her incredible support and guidance throughout the learning-and-writing process of this thesis project, as well as introducing me to the subject of this thesis in the first place. I would also like to thank Sejad Ali Kais and David Sermoneta for last-minute conversations concerning analyticity. Lastly, I would like to thank the community on Mathematics Stack Exchange for helping me with the computation of an integral, especially the fellow member with username “greg”.

Contents

Chapter 1

Introduction

The realization problem is a fundamental inverse problem in systems-and-control theory and concerns the construction of models for dynamical systems with data from its input-output behavior. The model is often called an *internal description* of the dynamical system and is typically given by a set of equations describing how the quantities of interest, as well as other measurable quantities, evolve over time. The input-output behavior is a so-called *external description* of the system as it specifies how the measurable quantities depend on how we influence the system, like a cause-and-effect relationship. For linear, time-invariant systems, it can, for example, be represented by a matrix-valued rational function called the *transfer function* which describes this relationship in the frequency domain.

The problem of realization for linear, time-invariant (finite-dimensional, deterministic) systems was solved in 1963 by Kalman and Gilbert in [Kal63] and [Gil63], respectively, using partial fraction decomposition of the transfer function to derive an internal description. This was only the beginning, as sub-problems of realization such as the *minimal* realization problem were still of interest. Great effort was put into understanding the problem and aspects such as dimension and minimality, which tells how simple or complex a model is, were connected to concepts such as observability and reachability, which are of great importance in designing feedback and observers.

A few years after his paper on the solution of the realization problem, Kalman posed the problem again, but now with a sequence of matrices derived from the transfer function, called a *Markov sequence*. Kalman solved the problem with Ho in [HK66], where an algorithm for finding a simplest possible internal description was presented. Kalman, and many others, later went on to work on the partial realization problem, in which we have an incomplete picture of what the external description of our system is. This was formulated by having only finitely many terms of a Markov sequence. The question of finding a simplest model in the case of an incomplete external description, was now of interest. It was revealed that solutions to this problem lose certain structure that make solutions to the minimal realization problem "nice" and "simple". Two notable works on the problem, aside from Kalman's, were Rissanen's work in 1971 ([Ris71]) dealing with recursive construction of models from partial input-output data, and Bosgra's in 1983 ([Bos83]) concerning the construction of a simplest model and the parameterization of all such models.

During the late 1980's and early 1990's a lot of effort was put into solving the general problem of matrix-valued rational interpolation, in which we want to find a matrix-valued rational function which interpolates some prescribed set of matrices. A connection was made between this problem and that of partial realization. As presented by Anderson & Antoulas in [AA90], one could actually solve the partial realization problem with the methods of rational interpolation by first performing an elementary transformation of the transfer function.

Thus, the problem of constructing simple models for our linear dynamical systems from input-output data was essentially solved, but practically it was not enough. In some modeling scenarios, our simplest possible models for a system would still be too complex for any analysis or control, so there was incentive to study the process of decomposing models to even smaller ones which only preserve, or even approximate, *some* properties of the original simplest model. This process was named *model-order reduction* and one such problem formulation is that of approximate realization. Approximate realization is a problem in which we want to find a model which approximates the external description, rather than exactly reproducing it, and a notable paper is [Kun78], which details an algorithm for constructing approximate realizations using singular value decomposition on a matrix made of the input-output data. This approach is used in many other areas of model-order reduction such as in balanced model reduction (see [Moo81]), and in noisy realization (see [JP85]).

In this thesis, we solve study these problems of realization, partial realization and approximate realization, and for each problem provide a few different solution methods.

The outline of the thesis is as follows. In Chapter 2, we present different ways to represent linear, time-invariant dynamical systems and how the realization problem can be stated with the transfer function and with the Markov sequence. The connection between transfer functions and Markov sequences is also established. We then cover basic realization theory from the view of Markov parameters, present solutions as well as study some of their properties. Lastly, we present the partial realization problem and detail in what ways it differs from the realization problem.

In Chapter 3, we solve the partial realization problem in the case where our transfer function is scalar-valued by re-interpreting it as a rational interpolation problem, and solving it by means of residue interpolation. After that, we solve the problem of finding the simplest model that reproduces an incomplete external description, and describe when such a model is essentially unique.

In Chapter 4, we formulate the approximate realization problem, and solve it by first solving a structured low-rank approximation problem, which in turn is solved by treating it as one of least squares. Here, we tie in some of the methods discussed in Chapter 3, since the presented model-order reduction method rests on finding a model for a constructed sequence that approximates the original Markov sequence we have.

We end the thesis with Chapter 5, where we shortly summarize what we have done, as well as bring up what else could be done from here.

Chapter 2

Realization and Partial Realization

The construction of models for dynamical systems from experimental input-output data is an important task in many disciplines of engineering and science. The problems of realization and partial realization can be posed as such tasks since we use what we know about the input-output behavior of the system to construct a model for that specific behavior.

In what follows, we introduce necessary system-theoretic concepts and terminology to formulate and solve the realization problem, as well as the problem of "minimal" realization. After that, we formulate the partial realization problem and point out key differences between it and the realization problem. This chapter is based largely on the chapters 6.5 and 6.6 on realization theory for linear systems in [Son98], as well as the report [Sch00].

2.1 State-space Models and Transfer Functions

2.1.1 Representation of Systems

Let $\mathbb{K}$ be either the real or the complex field, and $\tilde{\mathbb{K}}$ be the algebraic closure of $\mathbb{K}$. In this thesis, we work exclusively with linear, time-invariant, continuous-time dynamical systems, i.e., those governed by equations of the following type:

$$\Sigma : \begin{cases} \dot{x}(t) = Ax(t) + B\omega(t) \\ y(t) = Cx(t) + D\omega(t), \\ x(0) = \vec{0}, \end{cases} \tag{2.1.1}$$

where $A \in \mathbb{K}^{n\times n}$, $x(t) \in \mathbb{K}^n$, $\omega(t) \in \mathbb{K}^m$, and $y(t) \in \mathbb{K}^p$ for all admissible t. The matrices A, B, C, D are the **system-matrices** for the system Σ, but we will sometimes say that the matrix-tuple (A, B, C, D) itself is a system, in which case we write $\Sigma = (A, B, C, D)$.

The above representation is often called a **state-space model**, or **internal description**, of a dynamical system since we can solve for the state $x(t)$. Apart from this representation, we will also consider the **transfer-function model** which is a so-called an **external description** of the system since it reveals nothing about the state of the system. This description is given by the so-called **transfer function**. Transfer functions of systems governed by equations like Equation 2.1.1 are given by matrix-valued rational functions of the following form:

$$W(s) = D + C(sI - A)^{-1}B, \quad s \in \tilde{\mathbb{K}}.$$

To distinguish transfer functions of different dynamical systems Σ and Σ', we will denote them by W_Σ and $W_{\Sigma'}$. We need this since different systems can have the same transfer function. Consider the following example.

Example 1. Let T be non-singular. The systems $\Sigma_1 = (A, B, C, D)$ and $\Sigma_2 = (T^{-1}AT, T^{-1}B, CT, D)$ have the same transfer function.

The act of going from one representation to another is of interest since control problems in one might be easier to solve than in another. Going from internal to external involves a routine calculation and the resulting representation is unique. Going from external to internal, however, is highly nontrivial since many different systems share the same transfer function and there isn't really a "one true way" to go about it. This problem, going from external to internal, is what is usually referred to as *realization*.

2.1.2 Realization of Transfer Functions

The problem of realization can be stated in the following way: if W is a matrix-valued rational function, does there exist a system $\Sigma = (A, B, C, D)$ for which $W_\Sigma = W$? The answer to the proposed question is unsurprisingly: it depends on W. If it does exist we say that W is **realizable**, that Σ **realizes** W and that W_Σ (or Σ) is a **realization** of W.

Now one would naturally ask: what property must W have for it to be realizable? The answer to this question is given by the following result.

Proposition 2.1.1. The matrix-valued rational function W is realizable if and only if it is **proper**, that is, the (entry-wise) limit

$$W(\infty) := \lim_{|s|\to\infty} W(s)$$

exists and is finite. In other words, W with values in $\tilde{\mathbb{K}}^{p\times m}$ is realizable if and only if $W(\infty) \in \tilde{\mathbb{K}}^{p\times m}$.

Proof. [BGR88, Chapter 2.4]. □

Remark 2.1.2. Note for each $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ that $\tilde{\mathbb{K}} = \mathbb{C}$.

Remark 2.1.3. What the above theorem *really* says is that a matrix-valued rational function is a transfer function of *some* system if and only if it is proper. To be brief, we will say that W is a transfer function if (and only if) it is proper.

Example 2. In the case of scalar-valued rational functions, properness is equivalent to the degree of the numerator-polynomial of the rational function being no greater than the degree of the denominator-polynomial. Thus, all rational functions $f = p/q$ where $\deg q \geq \deg p$ are transfer functions.

Suppose now that we want to find a realization for W, which is a proper matrix-valued rational function with values in $\tilde{\mathbb{K}}^{p\times m}$. By the above result, we at least know that there is a (currently unknown) system $\Sigma = (A, B, C, D)$ such that

$$W(s) = W_\Sigma(s) = D + C(sI - A)^{-1}B.$$

A well-known fact is that $C(sI - A)^{-1}B$ is **strictly proper**, meaning $C(sI - A)^{-1}B \to \mathbf{0}$ as $|s| \to \infty$. This is because

$$C(sI - A)^{-1}B = \frac{C\operatorname{adj}(sI - A)B}{\det(sI - A)},$$

where the entries of $C\mathrm{adj}(sI - A)B$ are polynomial of degree at most $\deg\det(sI - A) - 1$. A consequence of this is that

$$D = W(\infty).$$

Hence, we only need to determine matrices A, B, C to find a realization of W since D can be calculated by taking a limit. Furthermore, D is unique since if $\Gamma = (A', B', C', D')$ also realizes W, then it must hold that

$$D' = W_\Gamma(\infty) = W(\infty) = W_\Sigma(\infty) = D.$$

Using Proposition A.1.1, we have that

$$W(s) = D + C(sI - A)^{-1}B = D + Cs^{-1}(I - s^{-1}A)^{-1}B,$$

which for $|s| > \|A\|_F$, means that

$$\begin{aligned} W(s) &= D + Cs^{-1}\left(\sum_{k=0}^{\infty} s^{-k}A^k\right)B \\ &= D + \sum_{k=0}^{\infty} s^{-k-1}(CA^kB) = D + \sum_{i=1}^{\infty} s^{-i}(CA^{i-1}B). \end{aligned}$$

Since W is proper, we may compare the above expansion to the entry-wise Laurent-expansion of W around infinity, given by

$$W(s) = D + \sum_{i=1}^{\infty} s^{-i}\mathcal{A}_i, \quad |s| \text{ sufficiently large},$$

where $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, ... \in \mathbb{K}^{p\times m}$ are called **Markov parameters** of W, and the sequence $\mathcal{A} = \{\mathcal{A}_i\}_{i=1}^{\infty}$ of Markov parameters is called the **Markov sequence** of W.

Since the Laurent expansion is unique with respect to where we expand, we have for sufficiently large $|s|$ that

$$\mathcal{A}_i = CA^{i-1}B, \quad \text{for all } i \in \mathbb{Z}^+.$$

This gives us another characterization of realization which is purely matrix-algebraic. In this characterization, the problem boils down to factoring a sequence of matrices in a certain way. This problem of matrix-factorization is also known as "the realization problem". We therefore state it below as such.

Problem 1 (Realization)**.** Given a Markov sequence $\mathcal{A} = \{\mathcal{A}_i\}_{i=1}^{\infty}$ of $p \times m$ matrices over the field $\mathbb{K}$, find $n \in \mathbb{Z}^+$ and matrices $A \in \mathbb{K}^{n\times n}, B \in \mathbb{K}^{n\times m}, C \in \mathbb{K}^{p\times n}$ such that

$$\mathcal{A}_i = CA^{i-1}B, \quad \text{for all } i \in \mathbb{Z}^+.$$

One would naturally ask the following question: if (A, B, C) solves Problem 1 for the Markov sequence of a transfer function W, is it really the case that

$$W = W_\Sigma$$

everywhere on the complex plane? Above we see for a sufficiently large L that the triple will for $|s| > L \geq \|A\|_F$ satisfy

$$W(s) = W(\infty) + \sum_{i=1}^{\infty} s^{-i}\mathcal{A}_i = W(\infty) + \sum_{i=1}^{\infty} s^{-i}(CA^{i-1}B) = W(\infty) + C(sI - A)^{-1}B,$$

but does not mention what happens for $|s| \leq L$. The answer is that the matrix-valued functions will also coincide there. To prove this, we use the uniqueness property of direct analytic continuations on each of the entries of the matrix-valued rational functions W and W_Σ.

Let f and g be the (i, j)th entry of W and W_Σ, respectively, and assume that they are irreducible, i.e., pole-zero cancellation has already been made. Let $L_0 \geq L$ be a constant such that the domain

$$D_0 := \{s \in \mathbb{C} : |s| > L_0\}$$

has no poles of either f and g, and consider also the domain

$$D_1 := \mathbb{C} \setminus \{\text{poles of } f \text{ or } g\}.$$

We have that $f = g$ on D_0. Thus, f is a direct analytic continuation of g to D_1. Similarly, g is a direct analytic continuation of f to D_1. By uniqueness, we have that $f = g$ on D_1. Now we consider the following, if g has a pole at z of order k, and f is not analytic at z, then f also must have a pole at z since it is an irreducible rational function. If f does not have a pole at z, then f is analytic at z, since, again, f is an irreducible rational function. Thus, for a sufficiently small $\epsilon > 0$, we have that

$$g(s) = (s - z)^{-k} h(s) \quad \text{for } 0 < |s - z| < \epsilon.$$

where h is analytic at z with $h(z) \neq 0$. Choose now $0 < \epsilon_0 < \epsilon$ such that any s satisfying $0 < |s - z| < \epsilon_0$ also is in D_1. We have then that

$$(s - z)^{-k} h(s) = f(s) \quad \text{for } 0 < |s - z| < \epsilon_0.$$

In particular, we have that $h(z) = (s - z)^k f(s)$. The right-hand-side is analytic at z, and thus is the unique direct analytic continuation of h to $\{s \in \mathbb{C} : |s - z| < \epsilon_0\}$. We have from this though that $h(z) = 0$, which is contradictory, since we assumed that $h(z) \neq 0$. Thus, either g has no pole in z or f is not analytic at z. Similar argument on the poles of f yield that the poles of the rational function coincide, meaning $f = g$ on $\mathbb{C}$.

2.2 Realization Theory for Linear, Time-invariant Systems

2.2.1 Existence of Realizations

A Markov sequence $\mathcal{A}$ of $p \times m$ matrices over $\mathbb{K}$ is realizable if there are matrices A, B, C that solve Problem 1 for $\mathcal{A}$. The triple $\Sigma = (A, B, C)$ is said to be a realization of $\mathcal{A}$. Note here the distinction between realizations of transfer functions and realizations of Markov sequences. The latter is only a matrix-triple but the former is a matrix 4-tuple. Despite this, we will use the symbol Σ for both types of realization. In this section, we refer to realizations of Markov parameters as "realizations".

Determining if a sequence has a realization boils down to studying the **rank** of the sequence. The rank of a sequence $\mathcal{A} \subset \mathbb{K}^{p \times m}$ is defined by

$$\operatorname{rank} \mathcal{A} := \sup_{s,t \in \mathbb{Z}^+} \operatorname{rank} \mathcal{H}_{s,t}(\mathcal{A})$$

where $\mathcal{H}_{s,t}(\mathcal{A})$ is the (s, t)th **block-Hankel matrix** of the sequence $\mathcal{A}$, given by

$$\begin{pmatrix} \mathcal{A}_1 & \mathcal{A}_2 & \cdots & \mathcal{A}_t \\ \mathcal{A}_2 & \mathcal{A}_3 & \cdots & \mathcal{A}_{t+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{A}_s & \mathcal{A}_{s+1} & \cdots & \mathcal{A}_{s+t-1} \end{pmatrix} \in \mathbb{K}^{ps \times mt}.$$

When the sequence is scalar-valued, its is instead just called the (s,t)th **Hankel matrix** of $\mathcal{A}$. In general, Hankel matrices are defined by having constant anti-diagonals.

Theorem 28 in [Son98, Chapter 6.5] states that a sequence is realizable if and only if its rank is finite. The calculation of the rank of a generic sequence is no trivial task, and thus determining realizability of any chosen sequence of $p \times m$ matrices is not easy in general. Not only is it difficult to determine realizability via the above criterion, it also doesn't reveal anything about *how* to obtain a realization. Therefore, one may want to characterize realizability in another way, like the following.

Theorem 2.2.1. *A Markov sequence $\mathcal{A}$ is realizable if and only if it is* ***recursive****, i.e., if there is a positive integer n and scalars $\alpha_1, \alpha_2 ..., \alpha_n \in \mathbb{K}$ such that*

$$\mathcal{A}_{n+k+1} = \alpha_n \mathcal{A}_{n+k} + \alpha_{n-1}\mathcal{A}_{n+k-1} + \cdots + \alpha_1 \mathcal{A}_{k+1}$$

for all $k \geq 0$. We say in this case that $\mathcal{A}$ is recursive for $(\alpha_1, \dots, \alpha_n)$.

Proof. Suppose (A, B, C) realizes $\mathcal{A}$. By the *Cayley-Hamilton Theorem*, we have that

$$A^n = \alpha_n A^{n-1} + \alpha_{n-1} A^{n-2} + \cdots + \alpha_1 I$$

where the coefficients come from

$$\chi_A(s) = \det(sI - A) = s^n - \alpha_n s^{n-1} - \cdots - \alpha_1.$$

By left-multiplication of CA^{k+2}, where $k \in \mathbb{N}$ is arbitrary, and right-multiplication by B, we get from the above equality that

$$\mathcal{A}_{n+k+1} = \alpha_n \mathcal{A}_{n+k} + \alpha_{n-1}\mathcal{A}_{n+k-1} + \cdots + \alpha_1 \mathcal{A}_{k+1},$$

meaning that $\mathcal{A}$ is recursive.
Suppose now instead that $\mathcal{A}$ is recursive for $(\alpha_1, \dots, \alpha_n)$. We set A to be the matrix

$$\begin{pmatrix} \mathbf{0} & I_p & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & I_p & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & I_p \\ \alpha_1 I_p & \alpha_2 I_p & \alpha_3 I_p & \dots & \alpha_n I_p \end{pmatrix} \in \mathbb{K}^{np \times np},$$

and denote $\mathcal{H}_{n+k,n}(\mathcal{A})$ with the first k block-rows, i.e., first kp rows, removed by $\mathcal{H}^k_{n+k,n}(\mathcal{A})$. Note for $k = 0$ that $\mathcal{H}^k_{n+k,n}(\mathcal{A}) = \mathcal{H}_{n,n}(\mathcal{A})$. Simple but tedious calculations show that

$$A\mathcal{H}^k_{n+k,n}(\mathcal{A}) = \mathcal{H}^{k+1}_{n+k+1,n}(\mathcal{A}),$$

and by induction, that

$$A^k \mathcal{H}_{n,n}(\mathcal{A}) = \mathcal{H}^k_{n+k,n}(\mathcal{A}) \text{ for all } k \in \mathbb{N}.$$

Thus, we have that

$$\mathcal{A}_k = \begin{pmatrix} I_p & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix} A^{k-1} \mathcal{H}_{n,n}(\mathcal{A}) \begin{pmatrix} I_m \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix}$$

$$= \begin{pmatrix} I_p & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix} A^{k-1} \begin{pmatrix} \mathcal{A}_1 \\ \mathcal{A}_2 \\ \mathcal{A}_3 \\ \vdots \\ \mathcal{A}_n \end{pmatrix}.$$

Setting

$$C = \begin{pmatrix} I_p & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} \mathcal{A}_1 \\ \mathcal{A}_2 \\ \mathcal{A}_3 \\ \vdots \\ \mathcal{A}_n \end{pmatrix},$$

we have found a realization for $\mathcal{A}$, given by (A, B, C). □

The constructed realization in the proof above is called the **observability-form** realization of $\mathcal{A}$ since it can be shown to be "observable".

Definition 2.2.2. A pair of matrices (A, C) where $A \in \mathbb{K}^{n\times n}$ and $C \in \mathbb{K}^{p\times n}$ is **observable** if its **observability matrix**

$$\mathbf{O}(A, C) := \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix}$$

is full-rank.

An alternative realization for a Markov sequence $\mathcal{A}$ which is recursive for $(\alpha_1, \dots, \alpha_n)$ is the so-called **controllability-form** realization, given by (F, G, H) where

$$F = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \alpha_1 I_m \\ I_m & \mathbf{0} & \cdots & \mathbf{0} & \alpha_2 I_m \\ \mathbf{0} & I_m & \cdots & \mathbf{0} & \alpha_3 I_m \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & I_m & \alpha_n I_m \end{pmatrix}$$

$$H = \begin{pmatrix} I_m \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix}$$

$$G = \begin{pmatrix} \mathcal{A}_1 & \mathcal{A}_2 & \mathcal{A}_3 & \cdots & \mathcal{A}_n \end{pmatrix}.$$

Unsurprisingly, its name comes from the fact that it is “controllable”.

Definition 2.2.3. A pair of matrices (A, B) where $A \in \mathbb{K}^{n\times n}$ and $B \in \mathbb{K}^{n\times m}$ is **controllable** if its **controllability matrix**

$$\mathbf{R}(A, B) \coloneqq \begin{pmatrix} B & AB & \cdots & A^{n-1}B \end{pmatrix}$$

is full-rank.

One thing to note is that determining a recursive relation for a sequence is also not trivial in general. But what the above characterization gives us is a clear way to realize our sequence provided we know it is recursive.

2.2.2 Minimal Realizations

As in Example 1, we see that realizations are not unique. We can for a recursive sequence construct both its observability-form and controllability-form realizations, which are different. It is then natural to ask how we choose a realization among these. One way to think about this is practically. We want a model for our dynamical system that is as simple as possible while also, obviously, having the exact same input-output behavior we have witnessed. The simplicity (or rather, complexity) of a realization $\Sigma = (A, B, C)$ is usually measured by its **dimension**, $\dim \Sigma$, which is given by the number of columns of the matrix A.

Thus, searching for a realization that is “as simple as possible” is equivalent to constructing a realization which has smallest dimension among all possible realizations, a so-called **minimal** realization. The minimal possible dimension a realization of a finite-rank Markov sequence is the rank of the sequence, according to Theorem 27 combined with Theorem 28 in [Son98, Chapter 6.5].

Minimality of a realization $\Sigma = (A, B, C)$ is, by Theorem 27 in [Son98, Chapter 6.5], equivalent to the condition of (A, B, C) being **canonical**, i.e., (A, B) is controllable and (A, C) is observable. To obtain a minimal realization of a Markov sequence, one can simply construct the observability-form realization and perform so-called *Kalman* decomposition (see Lemma 6.5.1 in [Son98, Chapter 6.5]) to obtain a canonical triple that realizes the Markov sequence. Alternatively, one can construct the controllability-form realization and then perform Kalman decomposition. The Kalman decomposition can be seen as performing two other decompositions in succession, the so-called controllability-and-observability decompositions (see Lemmas 3.3.3 in [Son98, Chapter 3.3] and the discussion on p.272 in [Son98, Chapter 6.2]). See example below.

Example 3 (Observability-form + Controllability Decomposition)**.** Suppose our Markov sequence is $\mathcal{A} = \mathbb{N}$. This sequence is realizable since $n = (n-1) + (n-2) - (n-3)$ for all naturals $n \geq 3$, meaning (by shift of indices) that

$$\mathcal{A}_{k+4} = \mathcal{A}_{k+3} + \mathcal{A}_{k+2} - \mathcal{A}_{k+1}, \quad \text{for all } k \in \mathbb{N},$$

where $\mathcal{A}_1 = 0, \mathcal{A}_2 = 1, \mathcal{A}_3 = 2$. The observability-form realization of $\mathcal{A}$ is given by

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 1 \end{pmatrix}, \; B = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \; C = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}.$$

Note that

$$\begin{aligned}\mathbf{R}(A,B) &= \begin{pmatrix} B & AB & A^2B \end{pmatrix}\\ &= \begin{pmatrix} 0 & 1 & 2\\ 1 & 2 & 3\\ 2 & 3 & 4 \end{pmatrix},\end{aligned}$$

which is does not have full-rank, meaning (A, B) is not controllable. By utilizing the `ControllableDecomposition`-function in *Wolfram Mathematica v14.3*, we find a decomposition of (A, B, C) which is controllable,

$$F = \begin{pmatrix} 1 & 0\\ -\sqrt{\frac{3}{2}} & 1 \end{pmatrix},\ G = \begin{pmatrix} -\sqrt{2}\\ \sqrt{3} \end{pmatrix},\ H = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

We further note that

$$\mathbf{O}(F,H) = \begin{pmatrix} H\\ FH \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}}\\ 0 & \frac{1}{\sqrt{3}} \end{pmatrix},$$

which has full rank, implying that (F, H) is observable. The realization (F, G, H) is therefore canonical, and thus also minimal.

2.2.3 Uniqueness of Realizations

In the previous section, we saw mentioned two similar methods for finding a minimal realization for a finite-rank Markov sequence. The methods do not necessarily produce the same realization. So one would naturally ask if there is a way all minimal realizations. The answer to this question is given by Theorem 27 in [Son98, Chapter 6.5], which states that that all minimal realizations are **similar**, i.e., that for any two minimal realizations $\Sigma_1 = (A, B, C)$ and $\Sigma_2 = (F, G, H)$ there is a non-singular matrix T, which we will call a **similarity transformation**, such that

$$\begin{aligned} F &= T^{-1}AT,\\ G &= T^{-1}B,\\ H &= CT. \end{aligned}$$

Thus, if we obtain one minimal realization we obtain all of them, via the parameterization

$$\mathcal{P}_{\Sigma_1}(T) = (T^{-1}AT, T^{-1}B, CT), \quad T \text{ non-singular}.$$

This similarity property can be used to prove a result which is used to derive a class of more direct minimal realization methods. Before presenting and proving the result, we need the following lemmas.

Lemma 2.2.4. A triple $\Sigma = (A, B, C)$ realizes a Markov sequence $\mathcal{A} = \{\mathcal{A}_i\}_{i=1}^{\infty}$ if and only if

$$\mathcal{H}_{s,t}(\mathcal{A}) = \mathbf{O}_s(A,C)\mathbf{R}_t(A,B)$$

for all $s, t \in \mathbb{Z}^+$, where

$$\mathbf{O}_s(A,C) := \begin{pmatrix} C\\ CA\\ \vdots\\ CA^{s-1} \end{pmatrix},\quad \mathbf{R}_t(A,B) := \begin{pmatrix} B & AB & \cdots & A^{t-1}B \end{pmatrix},$$

are so-called the $s-$observability matrix of (A, C) and the $t-$controllability matrix of (A, B), respectively. In particular, if $n = \dim \Sigma$, then

$$\mathcal{H}_{n,n}(\mathcal{A}) = \mathbf{O}(A, C)\mathbf{R}(A, B).$$

Proof. [Son98, Chapter 6.5]. □

Lemma 2.2.5. If $\Sigma = (A, B, C)$ is a canonical realization of the Markov sequence $\mathcal{A} = \{\mathcal{A}_i\}_{i=1}^{\infty}$, then

$$\operatorname{rank} \mathcal{H}_{s,t}(\mathcal{A}) = \dim \Sigma,$$

for all integers $s, t \geq \dim \Sigma$. In particular, if $n = \dim \Sigma$ is the rank of $\mathcal{A}$, then $\operatorname{rank} \mathcal{H}_{n,n}(\mathcal{A}) = n$.

Proof. [Son98, Chapter 6.5]. □

We now present the result.

Theorem 2.2.6. *Let $\mathcal{A}$ be a Markov sequence of rank $n < \infty$. If matrices $N \in \mathbb{K}^{np\times n}$ and $M \in \mathbb{K}^{n\times nm}$ are such that*

$$\mathcal{H}_{n,n}(\mathcal{A}) = NM \quad \textit{and} \quad \operatorname{rank} N = \operatorname{rank} M = n,$$

then

1. *there are matrices $A \in \mathbb{K}^{n\times n}$, $B \in \mathbb{K}^{n\times m}$ and $C \in \mathbb{K}^{p\times n}$ such that*

$$\mathbf{O}(A, C) = N \quad \textit{and} \quad \mathbf{R}(A, B) = M,$$

2. *and the matrices A, B, C, given by*

$$\begin{aligned} A &= N^{-L}\mathcal{H}^{1}_{n+1,n}(\mathcal{A})M^{-R}, \\ B &= N^{-L}\mathcal{H}_{n,n}(\mathcal{A})\begin{pmatrix} I_m \\ \mathbf{0} \end{pmatrix}, \\ C &= \begin{pmatrix} I_p & \mathbf{0} \end{pmatrix}\mathcal{H}_{n,n}(\mathcal{A})M^{-R}, \end{aligned}$$

where N^{-L} and M^{-R} are left-and-right pseudo-inverses of N and M respectively, form a minimal realization of $\mathcal{A}$.

Proof. To prove the first statement, note that NM is a full-rank factorization of $\mathcal{H}_{n,n}(\mathcal{A})$ by Lemma 2.2.5. Let now (F, G, H) be a minimal realization of $\mathcal{A}$. We then get another full-rank factorization

$$\mathcal{H}_{n,n}(\mathcal{A}) = \mathbf{O}(F, H)\mathbf{R}(F, G),$$

by Lemma 2.2.4. By Proposition A.1.4, there is a similarity matrix $S \in \mathrm{GL}_n(\mathbb{K})$ such that

$$\mathbf{O}(F, H) = NS \quad \text{and} \quad \mathbf{R}(F, G) = S^{-1}M.$$

Thus, we have that

$$\begin{aligned} N &= \mathbf{O}(F, H)S^{-1} = \begin{pmatrix} HS^{-1} \\ HFS^{-1} \\ \vdots \\ HF^{n-1}S^{-1} \end{pmatrix}, \\ M &= S\mathbf{R}(F, G) = \begin{pmatrix} SG & SFG & \cdots & SF^{n-1}G \end{pmatrix}. \end{aligned}$$

The matrices given by

$$\begin{aligned}A &= SFS^{-1} \in \mathbb{K}^{n\times n},\\ B &= SG \in \mathbb{K}^{n\times m},\\ C &= HS^{-1} \in \mathbb{K}^{p\times n}.\end{aligned}$$

satisfies the first statement. In addition, since (F, G, H) minimally realizes $\mathcal{A}$, (A, B, C) will as well since it is similar to (F, G, H), meaning

$$CA^{k-1}B = HS^{-1}(SFS^{-1})^{k-1}SG = HF^{k-1}G = \mathcal{A}_k \quad \text{for } i \in \mathbb{Z}^+.$$

We now prove that A, B, C can we written in a certain way. Since B is in the first m columns of M, we have that

$$B = M\begin{pmatrix} I_m \\ \mathbf{0} \end{pmatrix} = N^{-L}NM\begin{pmatrix} I_m \\ \mathbf{0} \end{pmatrix} = N^{-L}\mathcal{H}_{n,n}(\mathcal{A})\begin{pmatrix} I_m \\ \mathbf{0} \end{pmatrix},$$

since N^{-L} is well-defined by Proposition A.1.3. Similarly, since C is contained in the first p rows of N, we get that

$$C = \begin{pmatrix} I_p & \mathbf{0} \end{pmatrix}\mathcal{H}_{n,n}(\mathcal{A})M^{-R},$$

since M^{-R} is well-defined by Proposition A.1.3. We can get A by noting that

$$\mathcal{H}^1_{n+1,n}(\mathcal{A}) = NAM.$$

Thus, we have that

$$A = N^{-L}\mathcal{H}^1_{n+1,n}(\mathcal{A})M^{-R},$$

as expected. □

A slight generalization of the above theorem can be used to prove a well-known construction method of Ho & Kalman. The generalization is to instead use the matrix $\mathcal{H}_{s,t}(\mathcal{A})$ where $s, t \in \mathbb{Z}^+$ is large enough such that the rank of $\mathcal{H}_{s,t}(\mathcal{A})$ doesn't increase.

The following is a special case of the method that can be proved from Theorem 2.2.6. The proof is a direct application of the theorem and will therefore not be given here.

Corollary 1 (Special case of B. L. Ho)**.** Let $\mathcal{A}$ be Markov sequence of rank $n < \infty$ and $P \in \mathbb{K}^{pn\times pn}, Q \in \mathbb{K}^{mn\times mn}$ be non-singular such that

$$P\mathcal{H}_{n,n}(\mathcal{A})Q = \begin{pmatrix} I_n & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \in \mathbb{K}^{pn\times mn}.$$

Then (A, B, C) defined as

$$\begin{aligned}A &= \begin{pmatrix} I_n & \mathbf{0} \end{pmatrix} P\mathcal{H}^1_{n+1,n}(\mathcal{A})Q\begin{pmatrix} I_n \\ \mathbf{0} \end{pmatrix},\\ B &= \begin{pmatrix} I_n & \mathbf{0} \end{pmatrix} P\mathcal{H}_{n,n}(\mathcal{A})\begin{pmatrix} I_m \\ \mathbf{0} \end{pmatrix},\\ C &= \begin{pmatrix} I_p & \mathbf{0} \end{pmatrix}\mathcal{H}_{n,n}(\mathcal{A})Q\begin{pmatrix} I_n \\ \mathbf{0} \end{pmatrix}\end{aligned}$$

is a minimal realization of $\mathcal{A}$.

2.3 The Partial Realization Problem

Suppose we do not have access to the entire Markov sequence, but instead only the first $r > 0$ terms. We would still like to know if there is a transfer function having these terms as its first r Markov parameters, and how to construct it if it exists. This is what the *partial* realization problem is all about.

Formally, we deal with finite sequences of Markov parameters which we call truncated Markov sequences. We identify Markov sequences of r Markov parameters with points in $(\mathbb{K}^{p\times m})^r$, where the parameters themselves are $p \times m$ matrices over $\mathbb{K}$. In the case where $p = m = 1$, truncated Markov sequences will be identified by column-vectors in $\mathbb{K}^r$.

Continuing with the matrix-algebraic formulations, we present the partial realization problem as a factorization problem.

Problem 2 (Partial Realization)**.** Given a truncated Markov sequence $\mathcal{M} = (M_1, M_2, \ldots, M_r)$ of $p\times m$ matrices over the field $\mathbb{K}$, find $n \in \mathbb{Z}^+$ and matrices $A \in \mathbb{K}^{n\times n}, B \in \mathbb{K}^{n\times m}, C \in \mathbb{K}^{p\times n}$ such that

$$M_i = CA^{i-1}B, \quad \text{for all } i = 1, \ldots, r.$$

Like solutions to Problem 1, we say that a triple $\Sigma = (A, B, C)$ is a realization of truncated Markov sequence $\mathcal{M} = (M_1, ..., M_r) \in (\mathbb{K}^{p\times m})^r$, and that it realizes $\mathcal{M}$, if $M_i = CA^{i-1}B$ holds for all $i = 1, ..., r$. The definitions and notation of the concepts of dimension and canonicity of a realization remain the same.

To distinguish between realizations of Markov sequences and realizations of truncated Markov sequences, we sometimes may call the latter **partial realizations**.

2.3.1 Existence of Partial Realizations

Unlike realizations of (infinite) Markov sequences, realizations of truncated Markov sequences always exist. To show this, we embed our finite sequence in a realizable Markov sequence. In this way, we can utilize the solution methods for realization to solve the partial realization problem. Consider the following example.

Example 4. Let $\mathcal{M} = (M_1, \ldots, M_r) \in (\mathbb{K}^{p\times m})^r$ be a finite sequence. Since we do not care what any realization of $\mathcal{M}$ will give after the rth Markov parameter, we can extend $\mathcal{M}$ with a countable number of zero-matrices from $\mathbb{K}^{p\times m}$. Suppose $\mathcal{M}' = \{\mathcal{M}'\}_{i=1}^{\infty}$ is that sequence. Then we have for all $k \in \mathbb{N}$ that

$$\mathbf{0} = \mathcal{M'}_{r+k+1} = \alpha_r \mathcal{M'}_{r+k} + \alpha_{r-1}\mathcal{M'}_{r+k-1} + \cdots + \alpha_1 \mathcal{M'}_{k+1}$$

for $\alpha_1 = \alpha_2 = \cdots = \alpha_r = 0$ and $\mathcal{M}'_i = M_i$ for $i = 1, \ldots, r$. Thus, we can construct the observability-form realization of $\mathcal{M}'$ which is a partial realization of $\mathcal{M}$.

2.3.2 Minimal Partial Realizations

We have seen previously that minimality of a realization Σ is equivalent to it being canonical. This does not in general hold for partial realizations. We can therefore not reliably determine minimality by checking observability and controllability. This is demonstrated in the following example.

Example 5. Consider the truncated Markov sequence $\mathcal{M} = (1, 1, 2)^\top$. It can be realized by

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}.$$

This is a canonical realization since both

$$\mathbf{R}(A, B) = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 0 \end{pmatrix}, \text{ and } \mathbf{O}(A, C) = I_3$$

have full rank. But the triple (F, G, H) where

$$F = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, G = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, H = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

also realizes $\mathcal{M}$ and has a dimension lower than 3.

Remark 2.3.1. What is true, though, is that every minimal realization of a truncated Markov sequence is canonical, since otherwise we could decompose it via Kalman to a realization of smaller dimension, contradicting minimality.

2.3.3 Uniqueness of Partial Realizations

The similarity property among minimal realizations does not in general hold for minimal partial realizations. This means that we cannot parameterize all minimal partial realizations via a mapping

$$\mathcal{P}_\Sigma(T) = (T^{-1}AT, T^{-1}B, CT), \quad T \text{ non-singular}$$

for an arbitrary minimal partial realization $\Sigma = (A, B, C)$. As an example, consider the following.

Example 6. Consider again the truncated Markov sequence $\mathcal{M} = (1, 1, 2)^\top$. We have seen that a realization of this sequence is

$$F = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, G = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, H = \begin{pmatrix} 1 & 0 \end{pmatrix}.$$

Another one is given by

$$F' = \begin{pmatrix} 0 & 4 \\ 1 & -2 \end{pmatrix}, G' = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, H' = \begin{pmatrix} 1 & 1 \end{pmatrix}.$$

There is no realization of $\mathcal{M}$ of smaller dimension that 2. To see this, we assume the contrary. Thus, by the definition of dimension of a system, we must have a realization of dimension no lower that 1, i.e., there exists $A \in \mathbb{K}$, $B \in \mathbb{K}$ and $C \in \mathbb{K}$ such that

$$CB = 1, \quad CAB = 1, \quad CA^2B = 2.$$

Note from the first two equations that we have $A = 1$, but then $CA^2B = CB = 1 \neq 2$, so a realization of dimension 1 cannot exist for $\mathcal{M}$. This implies that (F, G, H) and (F', G', H') are minimal realizations of $\mathcal{M}$. Furthermore, they are not similar, i.e., there is no $T \in \mathrm{GL}_2(\mathbb{K})$ such that $(F', G', H') = (T^{-1}FT, T^{-1}G, HT)$. We show this now.

If there were a $T \in \mathrm{GL}_2(\mathbb{K})$ such that $(F', G', H') = (T^{-1}FT, T^{-1}G, HT)$. Then, clearly,

$$\det F' = \det(T^{-1}FT) = \det F.$$

By calculation, we find that $\det F = -1 \neq -4 = \det F'$. Hence the triples cannot be similar.

Another consequence of the non-similarity is that we cannot use Theorem 2.2.6 or Ho's algorithm in the way it was previously stated since the proof assumed similarity, as well as knowledge of the rank of a sequence, which is not applicable for finite sequences. To remedy this, we present the following.

Proposition 2.3.2. If $\mathcal{M} = (M_1, \ldots, M_{2r}) \in (\mathbb{K}^{p\times m})^{2r}$ is a finite sequence of matrices for which

$$\operatorname{rank} \mathcal{H}_{r,r}(\mathcal{M}) = \operatorname{rank} \mathcal{H}_{r+1,r}(\mathcal{M}) = \operatorname{rank} \mathcal{H}_{r,r+1}(\mathcal{M}) = r,$$

then there is a unique Markov sequence $\mathcal{M}'$ of rank r with the first $2r$ Markov parameters being

$$M_1, M_2, \ldots, M_{2r-1}, M_{2r}.$$

Proof. [Son98, Chapter 6.6]. □

The above result can, as in the case of the extension of a finite sequence by zeros, be seen as embedding the truncated Markov sequence in a realizable Markov sequence. The knowledge of the rank of the above Markov sequence is what allows us to construct a partial realization of our finite sequence by applying Ho's algorithm on the embedded Markov sequence.

Chapter 3

Interpolation and Minimal Realization

The finite nature of the partial realization problem makes it more difficult to construct minimal realizations but it also makes it easier for us to interpret and reformulate the problem in order to solve it. In this chapter, we will study one of these re-interpretations, as well as answer the question of how to find minimal realizations of truncated Markov sequences.

In the sequel, we build towards the treatment of partial realization as a problem of interpolation by scalar-valued transfer functions. After that, we solve the minimal partial realization problem, and show under what conditions the solutions are essentially unique.

3.1 Interpolation by Transfer Functions

In this section, we confine ourselves to the **single-input, single-output** (SISO) case of the partial realization problem in which our truncated Markov sequence is scalar. Furthermore, we assume that $\mathbb{K} = \mathbb{C}$. In this case, we can reformulate the problem to one of interpolation, and solve it by methods of so-called *residue interpolation*.

Systems which have scalar Markov parameters have scalar-valued transfer functions, and are called **SISO-systems**. An equivalent characterization for SISO-systems is that Σ is SISO if the matrices B, C are column-and-row vectors, respectively, and D is scalar. To distinguish SISO-systems from other systems, we will for the most part use lower-case letters for the matrices B, C, D.

3.1.1 Scalar-valued Transfer Functions

We have previously defined what a transfer function of a system is. In this section, we go a bit deeper and prove some properties concerning scalar-valued transfer functions. Certain properties discussed here are used to motivate the reformulation of the problem to one of interpolation.

Definition 3.1.1. The **McMillan-degree** of a scalar-valued transfer function $f = pq^{-1}$, where $p, q \in \mathbb{C}[s]$ are co-prime, is the degree of q. We denote it by $\deg_M f$ to distinguish it from the degree of a polynomial, denoted by $\deg(\cdot)$.

Remark 3.1.2. To make it easy for us, we only consider scalar-valued transfer functions that are non-constant.

The McMillan degree of a transfer function does not change if a scalar is added to the function. This is crucial for understanding that strictly proper and non-strictly proper transfer functions have the same complexity in terms of their McMillan degrees. Thus, we can chose to only study the properties of McMillan degree for strictly proper transfer functions.

Proposition 3.1.3. If $p, q \in \mathbb{C}[s]$ are co-prime, then so are $dq + p$ and q for every $d \in \mathbb{C}$.

Proof. Let $r = dq + p$. Clearly, 1 divides both q and r. Let t be a divisor for both q and r. Then it must divide $p = r - dq$, meaning that t divides $\gcd(p, q) = 1$. Thus, $\gcd(r, q) = 1$. □

Corollary 2. if $f = pq^{-1}$ is a scalar-valued transfer function where $p, q \in \mathbb{C}[s]$ are co-prime, then for every $d \in \mathbb{C}$,

1. $g = d + f$ is a transfer function, and
2. $\deg_M(f) = \deg_M(g)$.

Proof. Since f is proper, then so is g, which is equivalent to g being a transfer function by Proposition 2.1.1 and its remark. Furthermore, $g = d + pq^{-1} = (dq + p)q^{-1}$. Since $dq + p$ and q are co-prime by Proposition 3.1.3, we have that

$$\deg_M f = \deg_M(pq^{-1}) = \deg q = \deg_M((dq + p)q^{-1}) = \deg_M(d + pq^{-1}) = \deg_M g.$$

□

The following theorem gives a transfer-function description for what it means for a triple to be canonical.

Theorem 3.1.4. *Let f be a strictly proper transfer function with realization $(A, b, c, 0)$. The triple (A, b, c) is canonical if and only if $c\,\mathrm{adj}(sI - A)b$ and $\det(sI - A)$ are co-prime.*

Proof. Suppose f is realized by a non-canonical triple (A, b, c) together with $d = 0$. Then there is a canonical triple (F, g, h) that , with $d = 0$, realizes f, i.e.,

$$f(s) = \frac{c\,\mathrm{adj}(sI - A)b}{\det(sI - A)} = \frac{h\,\mathrm{adj}(sI - F)g}{\det(sI - F)}.$$

Since $\deg\det(sI - F) < \deg\det(sI - A)$, $c\,\mathrm{adj}(sI - A)b$ and $\det(sI - A)$ cannot be co-prime.
Suppose now instead that f is realized by $(A, b, c, 0)$ where $c\,\mathrm{adj}(sI - A)b$ and $\det(sI - A)$ are not co-prime. Then, by cancellation of common zeros, we have that

$$f(s) = \frac{c\,\mathrm{adj}(sI - A)b}{\det(sI - A)} = \frac{p(s)}{q(s)},$$

with $\deg p < \deg c\,\mathrm{adj}(sI - A)b$, $\deg q < \deg\det(sI - A)$ and p, q co-prime. Let p and q be given by

$$q(s) = q_0 + q_0 s + q_2 s^2 + \cdots + q_m s^m$$
$$p(s) = p_0 + p_1 s + p_2 s^2 + \cdots + p_n s^n + 0 \cdot s^{n+1} + \cdots + 0 \cdot s^{m-1}.$$

where $m = \deg q$ and $n = \deg p$. For simplicity's sake, we assume $q_m = 1$. Note that $n < m$ since f is strictly proper. In [nd], the triple (A', b', c') together with $d = 0$, given by

$$A' = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 1 \\ -q_{m-1} & -q_{m-2} & -q_{m-3} & \cdots & -q_0 \end{pmatrix} \in \mathbb{C}^{m\times m},$$

$$b' = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{C}^{m\times 1},$$

$$c' = \begin{pmatrix} p_0 & p_1 & \cdots & p_n & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{C}^{1\times m}.$$

is said to form a realization for f, which means that (A, b, c) cannot be minimal and thus, not canonical. We show now that this is indeed the case. Note firstly that $\det(sI - A') = q(s)$ by Proposition A.1.5. Thus, if we show that $c'\mathrm{adj}(sI - A')b' = p(s)$, we will be done since it follows directly that $f(s) = c'(sI - A')^{-1}b'$. We have that

$$(sI - A') = \begin{pmatrix} s & -1 & 0 & \cdots & 0 \\ 0 & s & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & s & -1 \\ q_{m-1} & q_{m-2} & q_{m-3} & \cdots & s + q_0 \end{pmatrix} \in \mathbb{C}^{m\times m},$$

and since $\mathrm{adj}(sI - A')b'$ gives the last column of the transpose of the co-factor matrix of A', we have that

$$\mathrm{adj}(sI - A')b' = \begin{pmatrix} c_{m1} \\ c_{m2} \\ \vdots \\ c_{mm} \end{pmatrix},$$

where $(c_{m1}, c_{m2}, \ldots, c_{mm})^\top$ is the last row of the co-factor matrix of $(sI - A')$. The first three entries are given by

$$c_{m1} = (-1)^{m+1} \begin{vmatrix} -1 & 0 & \cdots & 0 \\ s & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & s & -1 \end{vmatrix} = (-1)^{m+1}(-1)^{m-1} = 1,$$

$$c_{m2} = (-1)^{m+2} \begin{vmatrix} s & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & s & -1 \end{vmatrix} = (-1)^{m+2}(-1)^{m-2}s = s,$$

$$c_{m3} = (-1)^{m+3} \begin{vmatrix} s & 0 & \cdots & 0 \\ 0 & s & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & s & -1 \end{vmatrix} = (-1)^{m+3}(-1)^{m-3}s^2 = s^2,$$

and by calculation of the rest, we see that we can express the co-factors as

$$c_{im} = s^{i-1}, \quad 1 \leq i \leq m.$$

Thus, we have that

$$\operatorname{adj}(sI - A')b' = \begin{pmatrix} 1 \\ s \\ s^2 \\ \vdots \\ s^{m-1} \end{pmatrix},$$

meaning

$$c' \operatorname{adj}(sI - A')b' = \begin{pmatrix} p_0 \\ p_1 \\ \vdots \\ p_n \\ 0 \\ \vdots \\ 0 \end{pmatrix}^{\top} \begin{pmatrix} 1 \\ s \\ s^2 \\ \vdots \\ s^{m-1} \end{pmatrix} = p_0 + p_1 s + p_2 s^2 + \cdots p_n s^n = p(s).$$

□

Remark 3.1.5. The above constructed realization is sometimes called the **standard reachable** realization of f.

An immediate consequence is the following.

Corollary 3. If (A, b, c) is canonical and $f(s) = c(sI - A)^{-1}b$ then the McMillan-degree of f is given by the number of eigenvalues of A, counting multiplicities. Equivalently $\deg_M f$ is given by the number of columns of A.

Proof. Since (A, b, c) is canonical, we know that the numerator, $c\text{adj}(sI - A)b$, of f and the denominator, $\det(sI - A)$, are co-prime by Theorem 3.1.4. Thus, $\deg_M f = \deg \det(sI - A)$. □

What the above two results hints at is that the dimension of a canonical system is exactly the McMillan degree of its transfer function. We state it without proof as a corollary.

Corollary 4. The minimal dimension of a realization of a scalar-valued transfer function f is $\deg_M f$.

We now want to study what happens to the McMillan-degree of $f = p/q$ if we first compose it with $s \mapsto s^{-1}$. To be precise, we ask if the McMillan-degree of $g = f \circ (\cdot)^{-1}$ is ever the same as the McMillan-degree of f. The following lemma says that it is, under certain conditions.

Lemma 3.1.6. If $p, q \in \mathbb{C}[s]$ satisfy the following:

1. p, q are non-zero and co-prime,
2. $\deg q \geq \deg p$,
3. $q(0) \neq 0$,

then $a, b \in \mathbb{C}[s]$ defined by

$$a(s) = s^n p(s^{-1})$$
$$b(s) = s^n q(s^{-1}),$$

with $n = \deg q$, are co-prime.

Proof. Since p, q are co-prime, they share no roots. The non-zero roots of a and b are the reciprocals of the roots of p, q. So a and b cannot have any non-zero roots in common. If p has 0 as a root, then a will still have that zero, but b cannot have a zero root. Thus, a and b must be co-prime. □

The above lemma can be used to prove under the same conditions that f and $g = f \circ (\cdot)^{-1}$ have the same McMillan degree.

Theorem 3.1.7. *If f is a proper rational function with no pole in $s = 0$, then $g = f \circ (\cdot)^{-1}$ is also proper. Furthermore, $\deg_M f = \deg_M g$.*

Proof. Since $f \neq 0$ is proper, we have that

$$f(s) = \frac{p_0 + p_1 s + \cdots p_m s^m}{q_0 + q_1 s + \cdots + q_n s^n} = \frac{p(s)}{q(s)}$$

where $q_0 \neq 0$ and $m \leq n$ since f is proper and has no pole in $s = 0$. Thus, we see that

$$\begin{aligned} g(s) = f(s^{-1}) &= \frac{p_0 + p_1 s^{-1} + \cdots p_m s^{-m}}{q_0 + q_1^{-s} + \cdots + q_n s^{-n}} \\ &= \frac{p_0 + p_1 s^{-1} + \cdots p_m s^{-m}}{q_0 + q_1^{-s} + \cdots + q_n s^{-n}} \cdot \frac{s^n}{s^n} \\ &= \frac{a_0 s^n + a_1 s^{n-1} + \cdots a_m s^{n-m}}{b_0 s^n + b_1 s^{n-1} + \cdots + b_n}, \end{aligned}$$

meaning g is proper. If the numerator and denominator of f, as represented above, are co-prime, then so are the denominator and numerator of g by Lemma 3.1.6. This is because

$$g(s) = \frac{p_0 + p_1 s^{-1} + \cdots p_m s^{-m}}{q_0 + q_1^{-s} + \cdots + q_n s^{-n}} \cdot \frac{s^n}{s^n} = \frac{s^n p(s^{-1})}{s^n q(s^{-1})},$$

$q(0) = q_0 \neq 0$ and $\deg q = n \geq m = \deg p$. Thus, we have that $\deg_M f = \deg_M g$. $\square$

We now want to know the following: If (A, b, c, d) realizes f. Then what will realize $g = f \circ (\cdot)^{-1}$? The following gives a realization that works for points far away from the origin.

Proposition 3.1.8. If f is a transfer function with realization

$$f(s) = d + c(sI - A)^{-1} b,$$

and $0 \notin \sigma(A)$, then $g = f \circ (\cdot)^{-1}$ is given by

$$g(s) = d - cA^{-1}b - cA^{-2}(sI - A^{-1})^{-1} b.$$

for all s such that $|s| > \|A^{-1}\|_F$.

Proof. Since 0 is not an eigenvalue of A, A has a trivial nullspace, meaning it is non-singular. Thus, via Proposition A.1.1, the following computation works.

$$\begin{aligned}
g(s) = f(s^{-1}) &= d + c(s^{-1}I - A)^{-1} b \\
&= d + cs(I - sA)^{-1} b \\
&= d - csA^{-1}(sI - A^{-1})^{-1} b \\
&= d - cA^{-1}s(sI - A^{-1})^{-1} b \\
&= d - cA^{-1}\left(s \sum_{i=0}^{\infty} A^{-i} s^{-i-1}\right) b \\
&= d - cA^{-1}\left(\sum_{i=0}^{\infty} A^{-i} s^{-i}\right) b \\
&= d - cA^{-1}\left(I + \sum_{i=1}^{\infty} A^{-i} s^{-i}\right) b \\
&= d - cA^{-1}\left(I + A^{-1}\sum_{i=1}^{\infty} A^{-i+1} s^{-i}\right) b \\
&= d - cA^{-1}(I + A^{-1}(sI - A^{-1})^{-1}) b \\
&= d - cA^{-1}b - cA^{-2}(sI - A^{-1})^{-1} b.
\end{aligned}$$

$\square$

The above result will be an important component for the method of realization we will later discuss. With it we get a realization of g for $|s| > \|A^{-1}\|_F$. For our purposes, we we only need to know how the function looks for large $|s|$, so the fact that we don't have a realization for g for $|s| \leq \|A^{-1}\|_F$ is not a problem. Although, there is something to be said about it, but we leave that for to the reader to figure out.

3.1.2 Residues of the Resolvent Matrix

We intend to reformulate the partial realization problem as a problem of transfer-function interpolation. This formulation involves residues of matrix-valued functions, so we need to define what it means to integrate such functions in general and how to compute the integral for two special matrix-valued functions in particular.

Definition 3.1.9. By a **(complex) matrix-valued function** we mean a function $F\colon D \to \mathbb{C}^{p\times m}$, where $D \subseteq \mathbb{C}$. We say that F is integrable along a contour $\gamma \subset D$ if all of its entries are integrable along γ and define

$$\int_\gamma F(s)\mathrm{d}s := \left(\int_\gamma F_{ij}(s)\mathrm{d}s\right)_{1\leq i\leq p, 1\leq j\leq m}.$$

In other words, the integral of F is defined entry-wise.

Definition 3.1.10. Let $A \in \mathbb{C}^{n\times n}$. The **resolvent** of A is the matrix-valued function

$$R_A(s) := (sI - A)^{-1} = \frac{\operatorname{adj}(sI - A)}{\det(sI - A)}.$$

The resolvent of a matrix is an important part of the interpolation formulation we are building towards. This is because they specify where the points we want to interpolate between are, and how the interpolation conditions look in general. Thus, we need to know how integration of it looks. See below example.

Example 7. We now consider for a matrix $A \in \mathbb{C}^{n\times n}$ its resolvent

$$R_A(s) = (sI - A)^{-1} = \frac{\operatorname{adj}(sI - A)}{\det(sI - A)}.$$

The entries of $\operatorname{adj}(sI - A)$ are polynomials, which are holomorphic, so for a simple contour γ that encloses no eigenvalues of A we have

$$\int_\gamma R_A(s)\mathrm{d}s = \mathbf{0},$$

by applying *Cauchy's Theorem* entry-wise. If, however, γ does enclose an eigenvalue, and lets for simplicity's sake assume it encloses all of $\sigma(A)$, then we need to analyze further.

Suppose $J = T^{-1}AT$ is given in Jordan-normal form. We have that

$$R_J(s) = (sI - J)^{-1} = (sT^{-1}T - T^{-1}AT)^{-1} = T^{-1}R_A(s)T.$$

If A has $p \leq n$ distinct eigenvalues, we can write J in the form

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_p \end{pmatrix},$$

where J_i is the $n_i \times n_i$ Jordan-block of the ith eigenvalue of A, provided some ordering of the p eigenvalues have been made. Note that $n_1 + n_2 + \cdots + n_p = n$.

We now explicitly calculate what R_J is. We have that

$$(R_J(s))^{-1} = sI - \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_p \end{pmatrix} = \begin{pmatrix} sI_1 - J_1 & & & \\ & sI_2 - J_2 & & \\ & & \ddots & \\ & & & sI_p - J_p \end{pmatrix},$$

where I_i is the $n_i \times n_i$ identity, for each $i \leq p$. By multiplication of block-matrices, it can be then shown that

$$R_J(s) = \begin{pmatrix} (sI_1 - J_1)^{-1} & & & \\ & (sI_2 - J_2)^{-1} & & \\ & & \ddots & \\ & & & (sI_p - J_p)^{-1} \end{pmatrix}.$$

Now let λ_i, where $i \leq p$, be an eigenvalue A. We have that

$$R_{J_i}(s) = (sI - (\lambda I + N_i))^{-1} = ((s - \lambda)I - N_i)^{-1},$$

where J_i denotes the Jordan-block in J associated with λ_i, and N_i is the $n_i \times n_i$ matrix with 1's on the super-diagonal and 0 everywhere else. For $|s - \lambda_i| > ||N_i||_F$, we have by Proposition A.1.1 that

$$R_{J_i}(s) = \sum_{k=0}^{\infty} (s - \lambda_i)^{-k-1} N_i^k = \sum_{k=0}^{n_i} (s - \lambda_i)^{-k-1} N_i^k,$$

where n_i is the multiplicity of λ_i, since N_i is nil-potent with index n_i. Note that the right-hand-side is a finite sum, and so the series converges for all $s - \lambda_i$, meaning the condition "$|s - \lambda_i| > ||N_i||_F$" is redundant.

Thus, if ς is a simple contour enclosing λ_i, we have that

$$\begin{aligned} \int_\varsigma R_{J_i}(s) \mathrm{d}s &= \int_\varsigma \sum_{k=0}^{n_i} (s - \lambda_i)^{-k-1} N_i^k \mathrm{d}s \\ &= \sum_{k=0}^{n_i} \left(\int_\varsigma \frac{1}{(s - \lambda_i)^{k+1}} \mathrm{d}s \right) N_i^k = \left(\int_\varsigma \frac{1}{(s - \lambda_i)^{0+1}} \mathrm{d}s \right) I_i = 2\pi i I_i, \end{aligned}$$

We have therefore on a contour γ enclosing $\sigma(A)$ that

$$\begin{aligned} \int_\gamma R_A(s) \mathrm{d}s &= T \left(\int_\gamma R_J(s) \mathrm{d}s \right) T^{-1} \\ &= T \begin{pmatrix} \int_\gamma (sI_1 - J_1)^{-1} \mathrm{d}s & & & \\ & \int_\gamma (sI_2 - J_2)^{-1} \mathrm{d}s & & \\ & & \ddots & \\ & & & \int_\gamma (sI_p - J_p)^{-1} \mathrm{d}s \end{pmatrix} T^{-1} \\ &= T(2\pi i I) T^{-1} = 2\pi i I. \end{aligned}$$

With the above computation, we can consider the integral of a little more intricate matrix-valued rational function. The following example will be used later on.

Example 8. Let $A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{m\times m}$ and $M \in \mathbb{C}^{n\times m}$. We now want to compute

$$\int_\gamma R_A(s) M R_B(s) \mathrm{d}s,$$

where γ is a simple contour enclosing $\sigma(A) \cup \sigma(B)$. To do so, we consider the block-matrix

$$K = \begin{pmatrix} A & \mathbf{0} \\ M & B \end{pmatrix}.$$

we have that

$$R_K(s) = (sI - K)^{-1} = \begin{pmatrix} sI - A & \mathbf{0} \\ -M & sI - B \end{pmatrix}^{-1} = \begin{pmatrix} (sI - A)^{-1} & \mathbf{0} \\ (sI - A)^{-1} M (sI - B)^{-1} & (sI - B)^{-1} \end{pmatrix},$$

by Proposition A.1.6. The spectrum of K is exactly $\sigma(A) \cup \sigma(B)$, which can be seen from Proposition A.1.7, because

$$\det(sI - K) = \det(sI - A) \det(sI - B).$$

So, if γ is a simple contour that encloses $\sigma(A) \cup \sigma(B)$, we have by Example 7 that

$$\int_\gamma (sI - K)^{-1} \mathrm{d}s = \begin{pmatrix} \int_\gamma (sI - A)^{-1} \mathrm{d}s & \mathbf{0} \\ \int_\gamma (sI - A)^{-1} M (sI - B)^{-1} \mathrm{d}s & \int_\gamma (sI - B)^{-1} \mathrm{d}s \end{pmatrix} = 2\pi i I,$$

meaning that

$$\int_\gamma (sI - A)^{-1} M (sI - B)^{-1} \mathrm{d}s = \mathbf{0}.$$

3.1.3 Residue Interpolation

In this section, we introduce a special case of the results on the interpolation by transfer functions that are treated in [GM10]. Namely, the formulations for the SISO-case. But before that, we present a more general problem of rational interpolation.

By a *(two-sided) residue interpolation* problem we mean a problem where we for chosen (data-)matrices $\mathcal{A}_1, \mathcal{A}_2, W_1, W_2$, H want to find a matrix-valued function Q such that

$$\frac{1}{2\pi i} \int_\gamma (sI - \mathcal{A}_1)^{-1} W_1 Q(s) W_2 (sI - \mathcal{A}_2)^{-1} \mathrm{d}s = H, \tag{RI}$$

where γ is a contour enclosing $\sigma(\mathcal{A}_1) \cup \sigma(\mathcal{A}_2)$.

If either of the pairs $(\mathcal{A}_1, W_1)$ or $(\mathcal{A}_1, W_1)$ aren't specified, we have a *one-sided* residue interpolation problem, which is to find Q satisfying

$$\frac{1}{2\pi i} \int_\gamma Q(s) W_2 (sI - \mathcal{A}_2)^{-1} \mathrm{d}s = H, \text{ or } \frac{1}{2\pi i} \int_\gamma (sI - \mathcal{A}_1)^{-1} W_1 Q(s) \mathrm{d}s = H.$$

Residue interpolation problems can be viewed as a generalization of many types of interpolation problems. Consider the following simple examples.

Example 9 (Interpolation)**.** Let Q be scalar-valued and that

$$\begin{aligned}\mathcal{A}_2 &= \operatorname{diag}(s_1, s_2, \dots, s_n) \in \mathbb{C}^{n\times n},\\ W_2 &= \begin{pmatrix}1 & 1 & \cdots & 1\end{pmatrix} \in \mathbb{C}^{1\times n},\\ H &= \begin{pmatrix}h_1 & h_2 & \cdots & h_n\end{pmatrix} \in \mathbb{C}^{1\times n},\end{aligned}$$

then RI is equivalent to the following interpolation conditions:

$$Q(s_i) = h_i, \quad i = 1, \dots, n.$$

Example 10 (Right-tangential Interpolation)**.** Let Q be scalar-valued and that

$$\begin{aligned}\mathcal{A}_2 &= \operatorname{diag}(s_1, s_2, \dots, s_n) \in \mathbb{C}^{n\times n},\\ W_2 &= \begin{pmatrix}w_1 & w_2 & \cdots & w_n\end{pmatrix} \in \mathbb{C}^{1\times n},\\ H &= \begin{pmatrix}h_1 & h_2 & \cdots & h_n\end{pmatrix} \in \mathbb{C}^{1\times n},\end{aligned}$$

then RI is equivalent to the following *tangential* interpolation conditions:

$$Q(s_i)w_i = h_i, \quad i = 1, \dots, n.$$

Example 11 (Single-point Hermite interpolation)**.** Let Q be scalar-valued and that

$$\begin{aligned}\mathcal{A}_2 &= \begin{pmatrix} s^* & 1 & & & \\ & s^* & 1 & & \\ & & s^* & 1 & \\ & & & \ddots & \ddots \\ & & & & s^* \end{pmatrix} \in \mathbb{C}^{n\times n},\\ W_2 &= \begin{pmatrix}1 & 0 & \cdots & 0\end{pmatrix} \in \mathbb{C}^{1\times n},\\ H &= \begin{pmatrix}h_1 & h_2 & \cdots & h_n\end{pmatrix} \in \mathbb{C}^{1\times n},\end{aligned}$$

then RI is equivalent to the following *Hermite* interpolation conditions:

$$Q^{(i)}(s^*) = h_i, \quad i = 0, \dots, n-1.$$

We now present the SISO-case of the main theorem in [GM10], retaining some of the more general notation used in the formulation. Furthermore, the proof shown here works for the $p \times m$-case.

Theorem 3.1.11. *Let*

1. $Q(s) = D + C(sI - A)^{-1}B$ *be a scalar-valued transfer function where* (A, B, C) *is canonical,*
2. $\mathcal{A}_1, \mathcal{A}_2$ *be square matrices in* $\mathbb{K}^{n_1\times n_1}$ *and* $\mathbb{K}^{n_2\times n_2}$ *respectively, such that their spectra do not intersect with the set of poles for* Q*,*
3. γ *be a simple contour that encloses* $\sigma(\mathcal{A}_2) \cup \sigma(\mathcal{A}_2)$ *but not any poles of* Q*,*
4. W_1, W_2 *be matrices in* $\mathbb{K}^{n_1\times 1}$ *and* $\mathbb{K}^{1\times n_2}$ *respectively,*

then for a given $H \in \mathbb{K}^{n_1 \times n_2}$, *Q satisfies*

$$\frac{1}{2\pi i}\int_\gamma (sI - \mathcal{A}_1)^{-1} W_1 Q(s) W_2 (sI - \mathcal{A}_2)^{-1} ds = H,$$

if and only if the solutions Y_1, Y_2 *to*

$$\begin{aligned} -\mathcal{A}_1 Y_1 + Y_1 A + W_1 C &= \mathbf{0} \\ A Y_2 - Y_2 \mathcal{A}_2 + B W_2 &= \mathbf{0} \end{aligned}$$

satisfy the condition

$$-Y_1 Y_2 = H.$$

In this case, defining

$$\begin{aligned} V_1 &\coloneqq Y_1 B + W_1 D \in \mathbb{K}^{n_1 \times 1}, \\ V_2 &\coloneqq C Y_2 + D W_2 \in \mathbb{K}^{1 \times n_2}, \end{aligned}$$

we get that

$$\begin{aligned} (sI - \mathcal{A}_1)^{-1}(W_1 Q(s) - V_1) &= -Y_1 (sI - A)^{-1} B, \\ (Q(s) W_2 - V_2)(sI - \mathcal{A}_2)^{-1} &= -C(sI - A)^{-1} Y_2. \end{aligned}$$

Proof. Since (A, B, C) is canonical, we have that the set of poles for Q is exactly $\sigma(A)$. Furthermore, the solutions Y_1, Y_2 for the Sylvester equations exist by Proposition A.1.8 and are unique since the spectra of $\mathcal{A}_1$ and $\mathcal{A}_2$ are disjoint with the spectrum of A. Using $V_2 = CY_2 + DW_2$, we have that

$$\begin{aligned} [Q(s)W_2 - V_2](sI - \mathcal{A}_2)^{-1} &= \\ [(D + C(sI - A)^{-1}B)W_2 - V_2](sI - \mathcal{A}_2)^{-1} &= \\ [DW_2 - V_2](sI - \mathcal{A}_2)^{-1} + C(sI - A)^{-1} B W_2 (sI - \mathcal{A}_2)^{-1}&. \end{aligned}$$

Using $BW_2 = -AY_2 + Y_2\mathcal{A}_2$, we have get

$$\begin{aligned} [DW_2 - V_2](sI - \mathcal{A}_2)^{-1} + C(sI - A)^{-1} B W_2 (sI - \mathcal{A}_2)^{-1} &= \\ [DW_2 - V_2](sI - \mathcal{A}_2)^{-1} + C(sI - A)^{-1}(-AY_2 + Y_2\mathcal{A}_2)(sI - \mathcal{A}_2)^{-1} &= \\ [DW_2 - V_2](sI - \mathcal{A}_2)^{-1} - C(sI - A)^{-1} Y_2 + CY_2 (sI - \mathcal{A}_2)^{-1}&. \end{aligned}$$

The last equality holds since $(sI - A)^{-1}(sI - A) = I$ and $(sI - \mathcal{A}_2)(sI - \mathcal{A}_2)^{-1} = I$, meaning

$$(sI - A)^{-1} A = s(sI - A)^{-1} - I \;\text{ and }\; \mathcal{A}_2 (sI - \mathcal{A}_2)^{-1} = s(sI - \mathcal{A}_2)^{-1} - I.$$

Since $DW_2 - V_2 = CY_2$, we have that

$$\begin{aligned} [DW_2 - V_2](sI - \mathcal{A}_2)^{-1} - C(sI - A)^{-1} Y_2 + CY_2 (sI - \mathcal{A}_2)^{-1} &= \\ -C(sI - A)^{-1} Y_2&. \end{aligned}$$

A similar computation using $W_1 C = \mathcal{A}_1 Y_1 - Y_1 A$ and $V_1 = Y_1 B + W_1 D$ yields $(sI - \mathcal{A}_1)^{-1}(W_1 Q(s) - V_1) = -Y_1(sI - A)^{-1}B$. Thus, we have

$$\frac{1}{2\pi i}\int_\gamma \left[(sI-\mathcal{A}_1)^{-1}W_1 Q(s) W_2 (sI-\mathcal{A}_2)^{-1}\right] \mathrm{d}s =$$
$$\frac{1}{2\pi i}\int_\gamma \left[(sI-\mathcal{A}_1)^{-1}W_1 [Q(s) W_2 - V_2](sI-\mathcal{A}_2)^{-1}\right] \mathrm{d}s =$$
$$-\frac{1}{2\pi i}\int_\gamma \left[(sI-\mathcal{A}_1)^{-1}W_1 C(sI-A)^{-1}Y_2\right] \mathrm{d}s,$$

since we by Example 8 have that

$$-\frac{1}{2\pi i}\int_\gamma \left[(sI-\mathcal{A}_1)^{-1}W_1 V_2 (sI-\mathcal{A}_2)^{-1}\right] \mathrm{d}s = \mathbf{0}.$$

Using $W_1 C = -\mathcal{A}_1 Y_1 + Y_1 A$, we obtain

$$-\frac{1}{2\pi i}\int_\gamma \left[(sI-\mathcal{A}_1)^{-1}W_1 C(sI-A)^{-1}Y_2\right] \mathrm{d}s =$$
$$-\frac{1}{2\pi i}\int_\gamma \left[(sI-\mathcal{A}_1)^{-1}(-\mathcal{A}_1 Y_1 + Y_1 A)(sI-A)^{-1}Y_2\right] \mathrm{d}s =$$
$$-\frac{1}{2\pi i}\int_\gamma \left[(sI-\mathcal{A}_1)^{-1}Y_1 Y_2 - Y_1(sI-A)^{-1}Y_2\right] \mathrm{d}s.$$

We have by Example 7 that

$$\frac{1}{2\pi i}\int_\gamma (sI-\mathcal{A}_1)^{-1}\mathrm{d}s = I,$$
$$\frac{1}{2\pi i}\int_\gamma (sI-A)^{-1}\mathrm{d}s = \mathbf{0},$$

since γ does not enclose any eigenvalue of A. Hence,

$$-\frac{1}{2\pi i}\int_\gamma \left[(sI-\mathcal{A}_1)^{-1}Y_1 Y_2 - Y_1(sI-A)^{-1}Y_2\right] \mathrm{d}s = -Y_1 Y_2 = H.$$

For a proof of the other direction, we refer the reader to [GM10]. □

Remark 3.1.12. The matrices V_1 and V_2 may seem unintuitive but for the case where $\mathcal{A}_1$ and $\mathcal{A}_2$ do not share any eigenvalues with A, it just gives two more interpolation conditions:

$$V_1 = \frac{1}{2\pi i}\int_{\gamma_1} (sI-\mathcal{A}_1)^{-1}W_1 Q(s)\mathrm{d}s,$$
$$V_2 = \frac{1}{2\pi i}\int_{\gamma_2} Q(s) W_2 (sI-\mathcal{A}_2)^{-1}\mathrm{d}s$$

where γ_i are simple contours that enclose $\sigma(\mathcal{A}_i)$, $i \in \{1, 2\}$, but not any poles of Q. In practice, we may also specify what V_1 and V_2 are since they can be used to construct interpolants.

As a consequence of the above, we get the following results, giving a lower bound of the McMillan-degree of proper (matrix-valued) rational functions that solve to RI and detailing how we can construct rational functions Q that solve RI and have lowest possible McMillan-degree.

Corollary 5. If Q is realized by (A, B, C, D) and solves RI, then $\deg_M Q \geq \operatorname{rank} H$ provided $\sigma(\mathcal{A}_1) \cup \sigma(\mathcal{A}_2)$ and $\sigma(A)$ are disjoint.

Proof. [GM10]. ☐

Corollary 6. Assume the same hypothesis as Theorem 3.1.11, that $Y_1 Y_2 = -H$ is a full-rank factorization and that $D \in \mathbb{C}$ is fixed. Then the rational functions Q_1 and Q_2, realized by $\Sigma_1 = (A_1, B_1, C_1, D)$ and $\Sigma_2 = (A_2, B_2, C_2, D)$ respectively, where

$$\begin{aligned} A_1 &= Y_1^{-L} \mathcal{A}_1 Y_1 - Y_1^{-L} W_1 C_2 \\ B_1 &= Y_1^{-L}(V_1 - W_1 D), \\ C_1 &= (V_2 - D W_2) Y_2^{-R}, \end{aligned}$$

and

$$\begin{aligned} A_2 &= Y_2 \mathcal{A}_2 Y_2^{-R} - B_1 W_2 Y_2^{-R} \\ B_2 &= Y_1^{-L}(V_1 - W_1 D), \\ C_2 &= (V_2 - D W_2) Y_2^{-R}, \end{aligned}$$

solve RI. Furthermore, $\dim \Sigma_1 = \dim \Sigma_2 = \operatorname{rank} H$.

Proof. From the equations

$$\begin{aligned} \mathbf{0} &= -\mathcal{A}_1 Y_1 + Y_1 A + W_1 C \\ V_1 &= Y_1 B + W_1 D, \\ V_2 &= C Y_2 + D W_2, \end{aligned}$$

we have that

$$\begin{aligned} A &= Y_1^{-L} \mathcal{A}_1 Y_1 - Y_1^{-L} W_1 C \\ B &= Y_1^{-L}(V_1 - W_1 D) \\ C &= (V_2 - D W_2) Y_2^{-R} \\ D &= D, \end{aligned}$$

realizes Q. Similarly, the other equations combined with $D = D$ give the other realization. We have that the number of rows of A is given by the number of column of Y_2^{-R}, which is $\operatorname{rank} H$ since Y_2 is full row-rank. Thus, for $\Sigma = (A, B, C, D)$, we have that $\dim \Sigma = \operatorname{rank} H$. Similar argument gives the same dimension for the other realization. ☐

Remark 3.1.13. Note that the above triples (A_1, B_1, C_1) and (A_2, B_2, C_2) have to be canonical. Otherwise, we could decompose them via Kalman to triples that together with D represent realizations of the same transfer functions but of smaller dimension than $\operatorname{rank} H$, which is not possible according to Corollary 5.

3.1.4 Application to Partial Realization

In this section, we aim to use the developed tools of the previous ones to solve the partial realization problem in the case where we have an odd number of Markov parameters. Partial realization can be viewed as an interpolation problem at the point at infinity. Since our interpolation formulation implicitly assumes that the points we interpolate between are finite, we want to perform some kind of transformation on the problem to make it into one where we interpolate at a finite point on the complex plane.

The idea is from [AMY91, Chapter 4.3.c] and is to do the following. For a Markov sequence $\mathcal{M} = (M_1, M_2, \ldots, M_{2r-1})^\top$ and a chosen M_0, we construct a proper rational function f given by

$$f(s) = M_0 + M_1 s + M_2 s^2 + \cdots + M_{2r-1} s^r + O(s^{2r}),$$

by solving a Hermite interpolation problem at $s = 0$. We then pre-compose the function with $s \mapsto s^{-1}$, which gives another proper rational function function $g = f \circ (\cdot)^{-1}$ of with the same McMillan-degree as f, by Theorem 3.1.7. The first $2r - 1$ Markov parameters of g are $M_1, M_2 \ldots, M_{2r-1}$, and $g(\infty) = M_0$. By utilizing the transfer-function interpolation methods in [GM10], we can find a realization for f, and thus also a "local" realization for g. This realization is a partial realization of $\mathcal{M}$.

To make this approach of partial realization concrete, we begin by specifying what data-matrices we should use for the characterization and methods in [GM10].

Theorem 3.1.14 (Single-point Hermite interpolation)**.** *Let $k \in \mathbb{Z}^+$ and $Q_0, Q_1, Q_2, \ldots, Q_{2k-1}$ and s_0 be points on the complex plane. Any transfer function $Q(s) = d + c(sI - A)^{-1}b$ with canonical (A, b, c) that satisfies $Q(s_0) = Q_0$ and RI for*

$$\mathcal{A}_1 = \begin{pmatrix} s_0 & & & \\ 1 & s_0 & & \\ & \ddots & \ddots & \\ & & 1 & s_0 \end{pmatrix} \in \mathbb{C}^{k\times k},$$

$$W_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{C}^{k\times 1},$$

$$\mathcal{A}_2 = \begin{pmatrix} s_0 & 1 & & \\ & s_0 & \ddots & \\ & & \ddots & 1 \\ & & & s_0 \end{pmatrix} \in \mathbb{C}^{k\times k},$$

$$W_2 = \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{C}^{1\times k},$$

$$H = \mathcal{H}_{k,k}(\mathcal{Q}),$$

where $\mathcal{Q} = (Q_1, Q_2, \ldots, Q_{2k-1})^\top$, also satisfies the Hermite interpolation condition

$$Q^{(i)}(s_0) = i! \cdot Q_i, \quad \textit{for all } i = 0, 1, \ldots, 2k-1.$$

Proof. We have that

$$\mathcal{A}_2 = s_0 I + N,$$

where N has super-diagonal with 1's and other entries being 0. Thus,

$$R_{\mathcal{A}_2}(s) = \sum_{j=0}^{k} \frac{1}{(s-s_0)^{-j-1}} N^j.$$

Calculation shows that N^j is a non-zero matrix and has 1's on the jth super-diagonal for $j \leq k-1$. Thus, we have that

$$R_{\mathcal{A}_2}(s) = \begin{pmatrix} (s-s_0)^{-1} & (s-s_0)^{-2} & \cdots & (s-s_0)^{-k} \\ & (s-s_0)^{-1} & \cdots & (s-s_0)^{-k+1} \\ & & \ddots & \vdots \\ & & & (s-s_0)^{-1} \end{pmatrix}.$$

Left-multiplication by W_2 then gives

$$\begin{aligned} W_2(sI-\mathcal{A}_2)^{-1} &= \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} (s-s_0)^{-1} & (s-s_0)^{-2} & \cdots & (s-s_0)^{-k} \\ & (s-s_0)^{-1} & \cdots & (s-s_0)^{-k+1} \\ & & \ddots & \vdots \\ & & & (s-s_0)^{-1}. \end{pmatrix} \\ &= \begin{pmatrix} (s-s_0)^{-1} & (s-s_0)^{-2} & \cdots & (s-s_0)^{-k} \end{pmatrix}. \end{aligned}$$

In a similar fashion, we have that

$$R_{\mathcal{A}_1}(s) = \begin{pmatrix} (s-s_0)^{-1} & & & \\ (s-s_0)^{-2} & (s-s_0)^{-1} & & \\ \vdots & \vdots & \ddots & \\ (s-s_0)^{-k} & (s-s_0)^{-k+1} & \cdots & (s-s_0)^{-1} \end{pmatrix},$$

meaning

$$(sI-\mathcal{A}_1)^{-1}W_1 = \begin{pmatrix} (s-s_0)^{-1} \\ (s-s_0)^{-2} \\ \vdots \\ (s-s_0)^{-k} \end{pmatrix}.$$

We have therefore that

$$\begin{aligned} R_{\mathcal{A}_1}(s)W_1Q(s)W_2R_{\mathcal{A}_2}(s) &= Q(s) \begin{pmatrix} (s-s_0)^{-1} \\ (s-s_0)^{-2} \\ \vdots \\ (s-s_0)^{-k} \end{pmatrix} \begin{pmatrix} (s-s_0)^{-1} \\ (s-s_0)^{-2} \\ \vdots \\ (s-s_0)^{-k} \end{pmatrix}^{\top} \\ &= Q(s) \begin{pmatrix} (s-s_0)^{-2} & (s-s_0)^{-3} & \cdots & (s-s_0)^{-k-1} \\ (s-s_0)^{-3} & (s-s_0)^{-4} & \cdots & (s-s_0)^{-k-2} \\ \vdots & \ddots & \vdots & \vdots \\ (s-s_0)^{-k-1} & (s-s_0)^{-k-2} & \cdots & (s-s_0)^{-2k} \end{pmatrix}. \end{aligned}$$

Thus, by the *Generalized Cauchy's Integral Formula* we get

$$\frac{1}{2\pi i}\int_\gamma (sI-\mathcal{A}_1)^{-1}W_1Q(s)W_2(sI-\mathcal{A}_2)^{-1}\mathrm{ds} =$$

$$\frac{1}{2\pi i}\int_\gamma \begin{pmatrix} Q(s)(s-s_0)^{-2} & Q(s)(s-s_0)^{-3} & \cdots & Q(s)(s-s_0)^{-k-1} \\ Q(s)(s-s_0)^{-3} & Q(s)(s-s_0)^{-4} & \cdots & Q(s)(s-s_0)^{-k-2} \\ \vdots & \ddots & \cdots & \vdots \\ Q(s)(s-s_0)^{-k-1} & Q(s)(s-s_0)^{-k-2} & \cdots & Q(s)(s-s_0)^{-2k} \end{pmatrix} \mathrm{ds} =$$

$$\begin{pmatrix} \frac{1}{1!}Q^{(1)}(s_0) & \frac{1}{2!}Q^{(2)}(s_0) & \cdots & \frac{1}{k!}Q^{(k)}(s_0) \\ \frac{1}{2!}Q^{(2)}(s_0) & \frac{1}{3!}Q^{(3)}(s_0) & \cdots & \frac{1}{(k+1)!}Q^{(k+1)}(s_0) \\ \vdots & \ddots & \cdots & \vdots \\ \frac{1}{k!}Q^{(k)}(s_0) & \frac{1}{(k+1)!}Q^{(k+1)}(s_0) & \cdots & \frac{1}{(2k-1)!}Q^{(2k-1)}(s_0) \end{pmatrix}.$$

By the hypothesis we have that

$$\begin{pmatrix} \frac{1}{1!}Q^{(1)}(s_0) & \frac{1}{2!}Q^{(2)}(s_0) & \cdots & \frac{1}{k!}Q^{(k)}(s_0) \\ \frac{1}{2!}Q^{(2)}(s_0) & \frac{1}{3!}Q^{(3)}(s_0) & \cdots & \frac{1}{(k+1)!}Q^{(k+1)}(s_0) \\ \vdots & \ddots & \cdots & \vdots \\ \frac{1}{k!}Q^{(k)}(s_0) & \frac{1}{(k+1)!}Q^{(k+1)}(s_0) & \cdots & \frac{1}{(2k-1)!}Q^{(2k-1)}(s_0) \end{pmatrix} =$$

$$\mathcal{H}_{r,r}(\mathcal{Q}) =$$

$$\begin{pmatrix} Q_1 & Q_2 & \cdots & Q_k \\ Q_2 & Q_3 & \cdots & Q_{k+1} \\ \vdots & \ddots & \cdots & \vdots \\ Q_k & Q_{k+1} & \cdots & Q_{2k-1} \end{pmatrix},$$

so the interpolation condition is satisfied. □

Corollary 7. If $Q(s) = d + c(sI-A)^{-1}b \neq 0$ satisfies the hypothesis of Theorem 3.1.14 for $s_0 = 0$ and $Q_0, Q_1, \ldots, Q_{2r-1} \in \mathbb{C}$, then R given by

$$R(s) = d - cA^{-1}b - cA^{-2}(sI - A^{-1})^{-1}b,$$

has its first $2r-1$ Markov parameters being $Q_1, \ldots, Q_{2r-1}$. Furthermore, $R(\infty) = Q_0$ and $\deg_M R = \deg_M Q$.

Proof. Since Q satisfies the hypothesis of Theorem 3.1.14 for $s_0 = 0$, we have that

$$Q^{(i)}(0) = i! \cdot Q_i \quad \text{for all } i = 0, 1, \ldots, 2k-1.$$

Thus, we have that

$$Q(s) = Q_0 + Q_1 s + Q_2 s^2 + \cdots + Q_{2k-1}s^{2k-1} + O(s^{2k}),$$

meaning that

$$R(s) = Q(s^{-1}) = Q_0 + Q_1 s^{-1} + Q_2 s^{-2} + \cdots + Q_{2k-1}s^{-2k+1} + O(s^{-2k}).$$

For $|s| > \|A^{-1}\|_F$, we by Proposition 3.1.8 that

$$R(s) = T(s) \coloneqq d - cA^{-1}b - cA^{-2}(sI - A^{-1})^{-1}b.$$

Similarly, we have by Proposition A.1.1 that

$$T(s) = d - cA^{-1}b + \sum_{i=1}^{\infty} s^{-k}(-cA^{-2})A^{-(k-1)}b = \sum_{i=0}^{2k-1} Q_i s^{-i} + O(s^{-2k}),$$

for $|s| > \|A^{-1}\|_F$, meaning that $Q_0 = d - cA^{-1}b$ and that

$$(-cA^{-2})(A^{-1})^{k-1}b = Q_k, \quad \text{for all } k = 1, 2, \dots, 2k-1.$$

By Theorem 3.1.6, we have that $\deg_M R = \deg_M Q$ since $\chi_A(0) \neq 0$, Q is non-zero and proper. □

Remark 3.1.15. To actually utilize the above result, we need to first construct a rational function $Q(s) = Q_0 + Q_1 s + Q_{2r-1}s^{2k-1} + O(s^{2k})$. We may use either of the constructions presented in Corollary 6. Note though that we need to specify what V_1 and V_2 are. An example that works is

$$V_1 = \begin{pmatrix} Q_0 \\ Q_1 \\ \vdots \\ Q_{r-1} \end{pmatrix} \in \mathbb{C}^{r\times 1}, \quad V_2 = V_1^\top.$$

We refer the reader to [GM10] for the proof.

What we have uncovered above is that if (A, b, c, d) is a system with canonical (A, b, c) for which the transfer function has $M_0, M_1, \dots, M_{2r-1}$ as its first Maclaurin coefficients, then the system given by $(A', b', c', d') = (A^{-1}, b, -cA^{-2}, d - cA^{-1}b)$ satisfies $M_0 = d - cA^{-1}b$ and has $M_1, \dots, M_{2r-1}$ as its first $2r - 1$ Markov parameters.

We end this section with an example that ties everything together and illustrates how the above material can be used to construct partial realizations.

Example 12. We have the truncated Markov sequence $\mathcal{M} = (2, 1, 3)^\top$ which we want to realize. In this case, we have that $r = 2$, and by setting

$$\begin{aligned} \mathcal{A}_1 &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \\ W_1 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ W_2 &= \begin{pmatrix} 1 & 0 \end{pmatrix} \\ H &= \mathcal{H}_{2,2}(\mathcal{M}) = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}, \end{aligned}$$

as well as specifying

$$D = 0, \quad V_1 = \begin{pmatrix} 0 \\ 2 \end{pmatrix} \quad \text{and} \quad V_2 = V_1^\top,$$

we can utilize Corollary 6 to construct a realization. We first need a full-rank factorization of $-H$. Note that H, and thus $-H$, is non-singular, so what we want are two non-singular matrices whose product is $-H$. One such pair is obtained by a LU-factorization of H, which is given below

$$H = \underbrace{\begin{pmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{pmatrix}}_{L} \overbrace{\begin{pmatrix} 2 & 1 \\ 0 & \frac{5}{2} \end{pmatrix}}^{U}.$$

Setting $Y_1 = -L$ and $Y_2 = U$, we have found a full-rank factorization $-H = Y_1 Y_2$. Note that both Y_1 and Y_2 are non-singular, so

$$Y_1^{-L} = Y_1^{-1} = \begin{pmatrix} -1 & -1 \\ \frac{1}{2} & 0 \end{pmatrix},$$
$$Y_2^{-R} = Y_2^{-1} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{5} \\ 0 & \frac{2}{5} \end{pmatrix}.$$

By following the first construction in Corollary 6, we have a triple $\Sigma_1 = (A, b, c)$ given by

$$c = V_2 Y_2^{-1} = \begin{pmatrix} 0 & \frac{4}{5} \end{pmatrix}$$
$$b = Y_1^{-1} V_1 = \begin{pmatrix} 0 \\ -2 \end{pmatrix}$$
$$A = Y_1^{-1} \mathcal{A}_1 Y_1 - Y_1^{-1} W_1 C = \begin{pmatrix} 0 & \frac{4}{5} \\ 1 & -\frac{2}{5} \end{pmatrix}.$$

We now define the triple $\Sigma_2 = (F, g, h)$ by

$$F = A^{-1} = \begin{pmatrix} \frac{1}{2} & 1 \\ \frac{5}{4} & 0 \end{pmatrix},$$
$$g = b$$
$$h = -cA^{-2} = \begin{pmatrix} -\frac{1}{2} & -1 \end{pmatrix},$$

and show that this is a realization of $\mathcal{M}$. We have that

$$hg = \begin{pmatrix} -\frac{1}{2} & -1 \end{pmatrix} \begin{pmatrix} 0 \\ -2 \end{pmatrix} = 2,$$
$$hFg = \begin{pmatrix} -\frac{1}{2} & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 1 \\ \frac{5}{4} & 0 \end{pmatrix} \begin{pmatrix} 0 \\ -2 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & -1 \end{pmatrix} \begin{pmatrix} -2 \\ 0 \end{pmatrix} = 1,$$
$$hF^2g = \begin{pmatrix} -\frac{1}{2} & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 1 \\ \frac{5}{4} & 0 \end{pmatrix} \begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & -1 \end{pmatrix} \begin{pmatrix} -1 \\ -\frac{5}{2} \end{pmatrix} = 3.$$

Furthermore, we have that $\dim \Sigma_2 = \dim \Sigma_1 = 2 = \operatorname{rank} H$, but also that Σ_2 is canonical since

$$\mathbf{O}(F, h) = \begin{pmatrix} h \\ hF \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & -1 \\ -\frac{3}{2} & -\frac{1}{2} \end{pmatrix}, \text{ and}$$
$$\mathbf{R}(F, g) = \begin{pmatrix} g & Fg \end{pmatrix} = \begin{pmatrix} 0 & -2 \\ -2 & 0 \end{pmatrix},$$

are both full-rank.

3.2 Minimal Partial Realizations

In this section we return back to the general case, where our Markov parameters are $p \times m$ matrices over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, to solve the minimal partial realization problem.

We have previously seen that canonicity of a partial realization is not sufficient for it to be minimal. Luckily for us, a very similar set of properties is. We present them below, and show how to obtain realizations satisfying them.

3.2.1 A Minimality Condition

Theorem 3.2.1. *A $q-$dimensional realization $\Sigma = (A, B, C)$ of the truncated Markov sequence $\mathcal{M} = (M_1, \ldots, M_r) \in (\mathbb{K}^{p\times m})^r$ is minimal if and only if the following hold:*

(i) $\ker \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{r-1} \end{pmatrix} = \{\vec{0}\}$,

(ii) $\operatorname{Im} \begin{pmatrix} B & AB & \cdots & A^{r-1}B \end{pmatrix} = \mathbb{K}^q$,

(iii) $\ker \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{j-1} \end{pmatrix} \subseteq \operatorname{Im} \begin{pmatrix} B & AB & \cdots & A^{r-j-1}B \end{pmatrix}$ *for all* $j = 1, \ldots, r-1$.

Notation 3.2.2. For the triple $\Sigma = (A, B, C)$ we will for $k \in \mathbb{Z}^+$ denote

$$\ker \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{pmatrix} = \bigcap_{j=1}^{k} \ker(CA^{j-1}) \text{ and,}$$

$$\operatorname{Im} \begin{pmatrix} B & AB & \cdots & A^{k-1}B \end{pmatrix} = \sum_{j=1}^{k} \operatorname{Im}(A^{k-1}B),$$

by $\mathcal{O}_k(\Sigma)$ and $\mathcal{R}_k(\Sigma)$ respectively. When there is no ambiguity about which triple the reachability subspaces or non-observability spaces are associated with, we write $\mathcal{O}_k$ and $\mathcal{R}_k$.

The approach in [GKL87] to proving the above minimality condition rests on a few lemmas which postulate the existence of decompositions that turn realizations into realizations that satisfy the properties $(i) - (iii)$. We now present these decompositions.

Lemma 3.2.3 (Observability-type Decomposition)**.** Let $\Sigma = (A, B, C)$ be a $q-$dimensional realization of $\mathcal{M} = (M_1, \ldots, M_r) \in (\mathbb{K}^{p\times m})^r$. There is a $T \in \mathrm{GL}_q(\mathbb{K})$ such that

$$F = T^{-1}AT = \begin{pmatrix} F_{10} & F_{01} \\ F_{11} & F_{00} \end{pmatrix},$$

$$G = T^{-1}B = \begin{pmatrix} G_1 \\ G_0 \end{pmatrix},$$

$$H = CT = \begin{pmatrix} \mathbf{0} & H_0 \end{pmatrix},$$

and the following hold:

1. $\Sigma_0 := (F_{00}, G_0, H_0)$ is a $(q - \dim \mathcal{O}_r(\Sigma))-$dimensional realization of $\mathcal{M}$,
2. $\mathcal{O}_r(\Sigma_0) = \{\vec{0}\}$.

Proof. Pick a subspace $\mathcal{Y}_0$ of $\mathbb{K}^q$ such that

$$\mathbb{K}^q = \mathcal{X}_1 \oplus \mathcal{Y}_0.$$

where $\mathcal{X}_1 = \mathcal{O}_r(\Sigma)$. Fix now a basis $X = \{x_1, \ldots, x_n\}$ for $\mathcal{X}_1$, where $n := \dim \mathcal{X}_1$, and a basis $Y = \{y_{n+1}, ..., y_q\}$ of $\mathcal{Y}_0$. We then define the matrices P_X and P_Y whose columns are made up of X and Y respectively, and lastly we define

$$T = \begin{pmatrix} P_X & P_Y \end{pmatrix} = \begin{pmatrix} x_1 & \cdots & x_n & y_{n+1} & \cdots & y_q \end{pmatrix} \in \mathrm{GL}_q(\mathbb{K}).$$

Consider now $H = CT$. Let $u \in \mathbb{K}^q$. Then there are $u_1 \in \mathbb{K}^n$ and $u_2 \in \mathbb{K}^{q-n}$ such that $u = (u_1^\top, u_2^\top)^\top$. Thus, we have that

$$CTu = C \begin{pmatrix} P_X & P_Y \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = CP_X u_1 + CP_Y u_2 = \vec{0} + CP_Y u_2,$$

since $\mathcal{X}_1 = \mathcal{O}_r(\Sigma) \subseteq \ker C$. Partitioning H into H_1 and H_0, where H_1 is made up of the first n columns of H, we get that

$$Hu = \begin{pmatrix} H_1 & H_0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = H_1 u_1 + H_0 u_2 = \vec{0} + CP_Y u_2.$$

Since this holds for all $u \in \mathbb{K}^q$, we have that $H_1 = \mathbf{0} \in \mathbb{K}^{p \times n}$.
Since $\mathcal{O}_r(\Sigma) \subseteq \ker(CA)$, we have that

$$HF = CAT = \begin{pmatrix} \mathbf{0} & H_0 \end{pmatrix} \begin{pmatrix} F_{10} & F_{01} \\ F_{11} & F_{00} \end{pmatrix} = \begin{pmatrix} H_0 F_{00} & H_0 F_{00} \end{pmatrix} = \begin{pmatrix} \mathbf{0} & H_0 F_{00} \end{pmatrix}.$$

By iterating the same argument for $k = 2, 3, ..., r-1$, we also get that

$$HF^k = CA^k T = \begin{pmatrix} \mathbf{0} & H_0 F_{00}^k \end{pmatrix}$$

since $\ker(CA^k) \supseteq \mathcal{O}_r(\Sigma)$ for the same k. Thus, we have clearly that

$$\begin{aligned} M_k &= CA^{k-1}B \\ &= (CT)(T^{-1}A^{k-1}T)(T^{-1}B) = (CT)(T^{-1}AT)^{k-1}(T^{-1}B) \\ &= HF^{k-1}G = \begin{pmatrix} \mathbf{0} & H_0 F_{00}^{k-1} \end{pmatrix} \begin{pmatrix} G_1 \\ G_0 \end{pmatrix} = H_0 F_{00}^{k-1} G_0, \quad \text{for } k = 1, ..., r, \end{aligned}$$

implying that $\Sigma_0 := (F_{00}, G_0, H_0)$ is a realization of $\mathcal{M}$. Its dimension is $q - n$ by construction, and since $n = \dim \mathcal{O}_r(\Sigma)$, we find that $\dim \Sigma_0 = q - \dim \mathcal{O}_r(\Sigma)$.
Consider now $x_0 \in \mathcal{O}_r(\Sigma_0) \subseteq \mathbb{K}^{q-n}$. Extending x_0 to $x = (\vec{0}^\top, x_0^\top)^\top \in \mathbb{K}^q$, we see firstly that $Tx \in \mathcal{Y}_0$ and secondly that

$$HF^k x = CA^k Tx = \begin{pmatrix} \mathbf{0} & H_0 F_{00}^k \end{pmatrix} \begin{pmatrix} \vec{0} \\ x_0 \end{pmatrix} = H_0 F_{00}^k x_0 = \vec{0}, \quad k = 0, \ldots, r-1,$$

so $Tx \in \mathcal{O}_r(\Sigma) = \mathcal{X}_1$. But since $\mathcal{X}_1$ and $\mathcal{Y}_0$ are direct complements, we must have that $Tx = \vec{0}$, meaning $x_0 = \vec{0} \in \mathbb{K}^{q-n}$. Thus, we must have that $\mathcal{O}_r(\Sigma_0) = \{\vec{0}\}$. ☐

Lemma 3.2.4 (Controllability-type Decomposition)**.** Let $\Sigma = (A, B, C)$ be a $q-$dimensional realization of $\mathcal{M} = (M_1, \ldots, M_r) \in (\mathbb{K}^{p\times m})^r$. There is a $T \in \mathrm{GL}_q(\mathbb{K})$ such that

$$F = T^{-1}AT = \begin{pmatrix} F_{10} & F_{01} \\ F_{11} & F_{00} \end{pmatrix},$$
$$G = T^{-1}B = \begin{pmatrix} \mathbf{0} \\ G_0 \end{pmatrix},$$
$$H = CT = \begin{pmatrix} H_1 & H_0 \end{pmatrix}$$

and the following hold:

1. $\Sigma_0 \coloneqq (F_{00}, G_0, H_0)$ is a $\dim \mathcal{R}_r(\Sigma)-$dimensional realization of $\mathcal{M}$,
2. $\mathcal{R}_r(\Sigma_0) = \mathbb{K}^{\dim \mathcal{R}_r(\Sigma)}$.

Proof. Pick a subspace $\mathcal{X}_1$ of $\mathbb{K}^q$ such that

$$\mathbb{K}^q = \mathcal{X}_1 \oplus \mathcal{Y}_0,$$

where $\mathcal{Y}_0 = \mathcal{R}_r(\Sigma)$. Fix now a basis $X = \{x_1, \ldots, x_n\}$ for $\mathcal{X}_1$, where $n \coloneqq \dim \mathcal{X}_1$, and a basis $Y = \{y_{n+1}, ..., y_q\}$ of $\mathcal{Y}_0$. We then define the matrices P_X and P_Y whose columns are made up of X and Y respectively, and lastly we define

$$T = \begin{pmatrix} P_X & P_Y \end{pmatrix} = \begin{pmatrix} x_1 & \cdots & x_n & y_{n+1} & \cdots & y_q \end{pmatrix} \in \mathrm{GL}_q(\mathbb{K}).$$

Now consider the matrix $G = T^{-1}B$. We have for all $u \in \mathbb{K}^m$ that

$$Bu = P_Y u'$$

for some $u' \in \mathbb{K}^{q-n}$ since $\mathrm{Im}(B) \subseteq \mathcal{R}_r(\Sigma) = \mathcal{Y}_0$. By setting $u'' = (\vec{0}^\top, u'^\top)^\top \in \mathbb{K}^q$, we get that

$$P_Y u' = \begin{pmatrix} P_X & P_Y \end{pmatrix} \begin{pmatrix} \vec{0} \\ u' \end{pmatrix} = Tu'',$$

so, it means that $Gu = T^{-1}(Tu'') = u''$. By partitioning G to $\begin{pmatrix} G_1^\top & G_2^\top \end{pmatrix}^\top$, where G_1 is the first n rows of G, we get for all $u \in \mathbb{K}^m$ that

$$Gu = \begin{pmatrix} G_1 \\ G_0 \end{pmatrix} u = \begin{pmatrix} G_1 u \\ G_0 u \end{pmatrix} = \begin{pmatrix} \vec{0} \\ u' \end{pmatrix},$$

so it must mean that $G_1 = \mathbf{0} \in \mathbb{K}^{n\times m}$. With this we have that

$$FG = \begin{pmatrix} F_{10} & F_{01} \\ F_{11} & F_{00} \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ G_0 \end{pmatrix} = \begin{pmatrix} F_{01}G_0 \\ F_{00}G_0 \end{pmatrix} = T^{-1}AB.$$

Since $\mathrm{Im}(AB) \subseteq \mathcal{R}_r(\Sigma)$, we have again that $F_{01}G_0 = \mathbf{0}$. So,

$$FG = \begin{pmatrix} \mathbf{0} \\ F_{00}G_0 \end{pmatrix}.$$

Using the same argument for $k = 2, 3, ..., r-1$, we also get that

$$F^k G = T^{-1} A^k B = \begin{pmatrix} \mathbf{0} \\ F_{00}^k G_0 \end{pmatrix},$$

since $\operatorname{Im}(A^k B) \subseteq \mathcal{R}_r(\Sigma)$ for the same k. Thus, we have clearly that

$$M_k = HF^{k-1}G = \begin{pmatrix} H_1 & H_0 \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ F_{00}^{k-1} G_0 \end{pmatrix} = H_0 F_{00}^{k-1} G_0, \quad k = 1, ..., r,$$

so $\Sigma_0 := (F_{00}, G_0, H_0)$ is a realization of $\mathcal{M}$.
To show that $q - n = \dim \mathcal{R}_r(\Sigma) = \dim \mathcal{Y}_0$, note that n is the dimension of $\mathcal{X}_1$ and that the direct sum of $\mathcal{X}_1$ and $\mathcal{Y}_0$ is $\mathbb{K}^q$ and has dimension q. Thus, $q - n = \dim \mathcal{Y}_0$.
We now show that $\mathcal{R}_r(\Sigma_0) = \mathbb{K}^{\dim \mathcal{R}_r(\Sigma)}$. By construction of Σ_0, we have that $\mathcal{R}_r(\Sigma_0) \subseteq \mathbb{K}^{q-n}$. Take now an arbitrary $x_0 \in \mathbb{K}^{q-n}$. By setting $x = (\vec{0}^\top, x_0^\top)^\top \in \mathbb{K}^q$, we note that

$$Tx = P_Y x_0 \in \mathcal{R}_r(\Sigma).$$

Thus, there are $u_k \in \mathbb{K}^m$ for $k = 0, \ldots r-1$ such that

$$\begin{aligned} T \begin{pmatrix} \vec{0} \\ x_0 \end{pmatrix} = Tx &= \sum_{k=0}^{r-1} A^k B u_k \\ &= \sum_{k=0}^{r-1} T F^k G u_k \\ &= \sum_{k=0}^{r-1} T \begin{pmatrix} \mathbf{0} \\ F_{00}^k G_0 \end{pmatrix} u_j = T \begin{pmatrix} \vec{0} \\ \sum_{k=0}^{r-1} F_{00}^k G_0 u_k \end{pmatrix}, \end{aligned}$$

which implies that $x_0 = \sum_{k=0}^{r-1} F_{00}^k G_0 u_j$ since T is one-to-one, meaning that $x_0 \in \mathcal{R}_r(\Sigma_0)$. □

Lemma 3.2.5 (Kalman-type Decomposition)**.** Let $\Sigma = (A, B, C)$ be a $q-$dimensional realization of $\mathcal{M} = (M_1, \ldots, M_r) \in (\mathbb{K}^{p \times m})^r$. Fix $j \in \{1, \ldots, r-1\}$. There is a $T \in \mathrm{GL}_q(\mathbb{K})$ dependent on j such that

$$\begin{aligned} F = T^{-1} A T &= \begin{pmatrix} F_{10} & F_{01} \\ F_{11} & F_{00} \end{pmatrix}, \\ G = T^{-1} B &= \begin{pmatrix} \mathbf{0} \\ G_0 \end{pmatrix}, \\ H = CT &= \begin{pmatrix} \mathbf{0} & H_0 \end{pmatrix} \end{aligned}$$

and the following holds:

1. $\Sigma_0 := (F_{00}, G_0, H_0)$ is a realization of $\mathcal{M}$ with
   $\dim \Sigma_0 = q - (\dim \mathcal{O}_j(\Sigma) - \dim\{\mathcal{O}_j(\Sigma) \cap \mathcal{R}_{r-j}(\Sigma)\})$,
2. $\mathcal{O}_j(\Sigma_0) \subseteq \mathcal{R}_{r-j}(\Sigma_0)$.

Proof. To prove this, fix $1 \leq j \leq r-1$ and pick a subspace $\mathcal{X}_1$ in $\mathcal{O}_j(\Sigma)$ such that

$$\mathcal{R}_{r-j}(\Sigma) + \mathcal{O}_j(\Sigma) = \mathcal{R}_{r-j}(\Sigma) \oplus \mathcal{X}_1.$$

Pick then a subspace $\mathcal{Y}_0 \supseteq \mathcal{R}_{r-j}(\Sigma)$ such that

$$\mathbb{K}^q = \mathcal{X}_1 \oplus \mathcal{Y}_0.$$

With the above decomposition in mind, we pick a basis $X = \{x_1, x_2, \ldots, x_n\}$ of $\mathcal{X}_1$ and a basis $Y = \{y_1, y_2, \ldots, y_{q-n}\}$ for $\mathcal{Y}_0$ such that $X \cup Y$ is a basis for $\mathbb{K}^q$. We then define the basis matrices P_X and P_Y whose columns are made up of X and Y respectively, and lastly we define

$$T = \begin{pmatrix} P_X & P_Y \end{pmatrix} = \begin{pmatrix} x_1 & \cdots & x_n & y_1 & \cdots & y_{q-n} \end{pmatrix} \in \mathrm{GL}_q(\mathbb{K}).$$

Consider now the matrix $G = T^{-1}B$. For any $u \in \mathbb{K}^m$, it is clear that $Gu = T^{-1}Bu$. Since $\mathcal{R}_{r-j}(\Sigma) \subseteq \mathcal{Y}_0$, we have in particular that $\mathrm{Im}(B) \subseteq \mathcal{Y}_0$. Thus, there is $u' \in \mathbb{K}^{q-n}$ such that $Bu = P_Y u'$. Extending u' by adding n zeroes, we get a $q-$vector $u'' = (\vec{0}^\top, u'^\top)^\top$ satisfying

$$Gu = T^{-1}Bu = T^{-1}\begin{pmatrix} P_X & P_Y \end{pmatrix}\begin{pmatrix} \vec{0} \\ u' \end{pmatrix} = T^{-1}Tu'' = u'' = \begin{pmatrix} \vec{0} \\ u' \end{pmatrix}.$$

Simultaneously, we also have that G can be split up to G_1 having the first n rows and G_0 having the rest, meaning that

$$Gu = \begin{pmatrix} G_1 \\ G_0 \end{pmatrix} u = \begin{pmatrix} G_1 u \\ G_0 u \end{pmatrix}.$$

But we know that any $m-$vector u multiplied by G gives a resulting vector $q-$vector with the first n coordinates being $\vec{0}$, so $G_1 = \mathbf{0}$.
In a similar way, we find that $H_1 = \mathbf{0}$.
The fact that Σ_1 is a realization of $\mathcal{M}$ is due to the following: (F, G, H) realizes $\mathcal{M}$ since it is similar to Σ. We have also via iteration on k that

$$F^k G = \begin{pmatrix} \mathbf{0} \\ F_{00}^k G_0 \end{pmatrix}, \quad \text{for } k = 0, 1, \ldots, r-j-1$$
$$HF^k = \begin{pmatrix} \mathbf{0} & H_0 F_{00}^k \end{pmatrix}, \quad \text{for } k = 0, 1, \ldots, j-1,$$

which together give

$$M_k = HF^{k-1}G = H_0 F_{00}^{k-1} G_0 \quad \text{for } k = 1, \ldots, r.$$

Now we need to prove that $\mathcal{O}_j(\Sigma_0) \subseteq \mathcal{R}_{r-j}(\Sigma_0)$, so we let $x_0 \in \mathcal{O}_j(\Sigma_0) \subseteq \mathbb{K}^{q-n}$ be non-zero. Then it means that

$$H_0 F_{00}^k x_0 = \vec{0}, \quad k = 0, ..., j-1$$

and by embedding x_0 in $\mathbb{K}^q$ as $x = (\vec{0}^\top, x_0^\top)^\top$, we have that

$$HF^k x = \begin{pmatrix} \mathbf{0} & H_0 F_{00}^k \end{pmatrix}\begin{pmatrix} \vec{0} \\ x_0 \end{pmatrix} = H_0 F_{00}^k x_0 = \vec{0}, \quad k = 0, ..., j-1.$$

But since $H = CT$ and $F = T^{-1}AT$, we get that

$$HF^k x = CT(T^{-1}AT)^k x = CA^k Tx = \vec{0}, \quad k = 0, ..., j-1.$$

meaning $Tx \in \mathcal{O}_j(\Sigma)$. Since $\mathcal{O}_j(\Sigma) \subseteq \mathcal{R}_{r-j}(\Sigma) \oplus \mathcal{X}_1$ and $\mathcal{R}_{r-j}(\Sigma) \subseteq \mathcal{Y}_0$, we have that $Tx \in \mathcal{R}_{r-j}(\Sigma)$ since, by definition of T, $Tx = P_Y x_0 \in \mathcal{Y}_0$ and $\mathcal{Y}_0 \cap \mathcal{X}_1 = \{\vec{0}\}$. If $x_0 = \vec{0} \in \mathbb{K}^{q-n}$ then $Tx = \vec{0}$ is still in $\mathcal{R}_{r-j}(\Sigma)$ since it is a subspace. So then we get that

$$\begin{aligned} T\begin{pmatrix} \vec{0} \\ x_0 \end{pmatrix} = Tx &= \sum_{k=0}^{r-j-1} A^k B u_k \\ &= \sum_{k=0}^{r-j-1} T F^k G u_k \\ &= \sum_{k=0}^{r-j-1} T \begin{pmatrix} \mathbf{0} \\ F_{00}^k G_0 \end{pmatrix} u_j = T \begin{pmatrix} \vec{0} \\ \sum_{k=0}^{r-j-1} F_{00}^k G_0 u_k \end{pmatrix} \end{aligned}$$

for some vectors $u_0, \ldots, u_{r-j-1} \in \mathbb{K}^m$. Since T is one-to-one, this means that

$$x_0 = \sum_{k=0}^{r-j-1} F_{00}^k G_0 u_k,$$

implying that $x_0 \in \mathcal{R}_{r-j}(\Sigma_0)$. Thus, $\mathcal{O}_j(\Sigma_0) \subseteq \mathcal{R}_{r-j}(\Sigma_0)$. To calculate the dimension of Σ_0, recall that

$$\dim \Sigma_0 = q - n = \dim \mathcal{Y}_0,$$

and that

$$\begin{aligned} n = \dim \mathcal{X}_1 &= \dim\{\mathcal{R}_{r-j}(\Sigma) + \mathcal{O}_j(\Sigma)\} - \dim \mathcal{R}_{r-j}(\Sigma) \\ &= (\dim \mathcal{R}_{r-j}(\Sigma) + \dim \mathcal{O}_j(\Sigma) - \dim\{\mathcal{R}_{r-j}(\Sigma) \cap \mathcal{O}_j(\Sigma)\}) - \dim \mathcal{R}_{r-j}(\Sigma) \\ &= \dim \mathcal{O}_j(\Sigma) - \dim\{\mathcal{R}_{r-j}(\Sigma) \cap \mathcal{O}_j(\Sigma)\}. \end{aligned}$$

Combining these yields the desired value of $\dim \Sigma_0$. □

Remark 3.2.6. The guaranteed existence of subspaces $\mathcal{X}_1$ and $\mathcal{Y}_0$ for a which, for a fixed $1 \leq j \leq r-1$, satisfy

1. $\mathcal{X}_1 \subseteq \mathcal{O}_j(\Sigma)$
2. $\mathcal{Y}_0 \supseteq \mathcal{R}_{r-j}(\Sigma)$
3. $\mathcal{X}_1 \oplus \mathcal{R}_{r-j}(\Sigma) = \mathcal{R}_{r-j}(\Sigma) + \mathcal{O}_j(\Sigma)$,
4. $\mathbb{K}^q = \mathcal{X}_1 \oplus \mathcal{Y}_0$,

is not entirely trivial. Here we give a proof. Let $B = \{v_1, \ldots, v_k\}$, where $k \leq q$, be a basis for the subspace $\mathcal{R}_{r-j}(\Sigma) + \mathcal{O}_j(\Sigma)$ of $\mathbb{K}^q$. Let $B_1 = \{v_{n_1}, v_{n_2}, \ldots, v_{n_p}\}$ be a subset of B consisting of vectors that are basis vectors for $\mathcal{O}_j(\Sigma)$ but not for $\mathcal{R}_{r-j}(\Sigma)$. Now set $\mathcal{X}_1 = \operatorname{span} B_1 \subseteq \mathcal{O}_j(\Sigma)$. Set $\mathcal{Y}_0$ to be the span of the remaining basis vectors of $\mathcal{R}_{r-j}(\Sigma) + \mathcal{O}_j(\Sigma)$ in addition to vectors needed to extend B to a basis for $\mathbb{K}^q$. This pair of subspaces satisfy the above conditions.

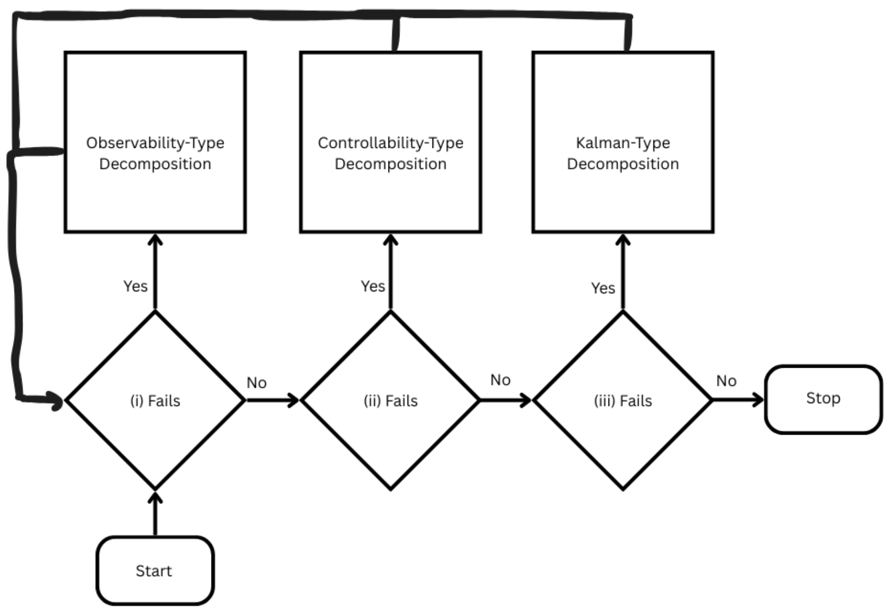
Note the fact that if Σ does not already satisfy the properties in Theorem 3.2.1, then there will be a guaranteed reduction of dimension when performing the decomposition according to Lemmas 3.2.3, 3.2.4, 3.2.5. This can be seen by the dimensions of the new realizations. For example, if $\mathcal{R}_r(\Sigma) \neq \mathbb{K}^{\dim \Sigma}$, meaning $\dim \mathcal{R}_r(\Sigma) < \dim \Sigma$, then decomposition via Lemma 3.2.4 will yield a realization Σ_0 with $\dim \Sigma_0 = \dim \mathcal{R}_r(\Sigma) < \dim \Sigma$.

3.2.2 The Compression Algorithm

The act of decomposing a realization according to Lemmas 3.2.4, 3.2.5, 3.2.3 depending on what property is missing is in [GKL87] referred to as a "compression algorithm" for finding a minimal partial realization. The fact that every realization can be decomposed to a realization satisfying $(i) - (iii)$ is not entirely trivial, and not explicitly proved. Here, we provide an argument for it.

The compression algorithm is not well described in [GKL87] since there is no mention of order of operations on the decompositions, i.e. no prescribed order on the properties. We shall here impose the order $(iii) \prec (ii) \prec (i)$, which means that property (i) should be checked and then property (ii) and lastly, property (iii) (See Figure 3.1). For the algorithm, this means that decompositions according to the above lemmas are made with respect to this order. Note that the proofs here should work with any other order.

Figure 3.1: Graphical representation of the compression algorithm with prescribed order



Proposition 3.2.7. If Σ_0 as in Lemmas 3.2.4, 3.2.3, or 3.2.5 has the same dimension as Σ then T is a similarity transformation.

Proof. If dimensions of the realizations are the same, then the sizes of the matrices F_{00}, F and A are the same, meaning that corresponding matrices are of the same size. □

Notation 3.2.8. When convenient, we will denote the image of a subspace $\mathcal{S}$ under a linear map T defined on $\mathcal{S}$ by $T\mathcal{S}$.

Proposition 3.2.9. Suppose $\Sigma = (A, B, C)$ is a $q-$dimensional realization of $\mathcal{M} = (M_1, \dots, M_r) \in (\mathbb{K}^{p\times m})^r$ satisfying the properties in Theorem 3.2.1 and $T \in \mathrm{GL}_q(\mathbb{K})$.
Then $\hat{\Sigma} = (T^{-1}AT, T^{-1}B, CT) = (F, G, H)$ is a realization of $\mathcal{M}$ satisfying the same properties in 3.2.1. In particular, properties $(i) - (iii)$ are preserved under similarity.

Proof. It is clear that $\hat{\Sigma}$ will realize $\mathcal{M}$, so we only prove the preservation of the properties. Note for $k \in \{1, \ldots, r\}$ that

$$\mathcal{O}_k(\Sigma) = \bigcap_{j=0}^{k-1} \ker(CA^j) = \bigcap_{j=0}^{k-1} \ker(HF^jT) = \bigcap_{j=0}^{k-1} T^{-1}\ker(HF^j) = T^{-1}\mathcal{O}_k(\hat{\Sigma}),$$

and that

$$\mathcal{R}_k(\Sigma) = \sum_{j=0}^{k-1} \mathrm{Im}(A^jB) = \sum_{j=0}^{k-1} \mathrm{Im}(T^{-1}F^jG) = T^{-1}\sum_{j=0}^{k-1} \mathrm{Im}(F^jG) = T^{-1}\mathcal{R}_k(\hat{\Sigma}),$$

The desired result follows from the above. □

Theorem 3.2.10. *The compression algorithm terminates with final state satisfying conditions* $(i) - (iii)$ *in Theorem 3.2.1.*

Proof. Let $\mathcal{M} = (M_1, \ldots, M_r)$ be a truncated Markov sequence and $\Sigma = (A, B, C)$ a realization of $\mathcal{M}$. We consider the sequence of realizations $\{\Sigma_i\}_{i\in\mathbb{N}}$, where Σ_i is the ith iteration of the compression algorithm, provided we start with $\Sigma_0 = \Sigma$. By construction, we have that

$$\dim \Sigma = \dim \Sigma_0 \geq \dim \Sigma_1 \geq \cdots \geq 1,$$

since the dimension is given by number of columns/rows of matrices, which are always at least 1. Thus, the sequence $\{\dim \Sigma_i\}_{i\in\mathbb{N}}$ forms a bounded subset of the natural numbers and must therefore have a minimum, $q \geq 1$. Suppose $n \in \mathbb{N}$ is an index for which

$$\dim \Sigma_n = q.$$

We have for all $k \in \mathbb{N}$ that

$$q = \dim \Sigma_n \geq \dim \Sigma_{n+k} \geq q,$$

So it must be that

$$\dim \Sigma_n = \dim \Sigma_{n+k}, \quad k \in \mathbb{N}.$$

We now show that any two realizations of $\{\Sigma_i\}_{i=n}^{\infty}$ are similar. Take Σ_k, Σ_l with $k > l \geq n$. If $k = l+1$, then by Proposition 3.2.7, they are similar. Similarly, for the case where $k > l+1$, we have that Σ_{l+p} and Σ_{l+p+1} are similar for $p = 0, 1, 2, ..., k-l-1$. Let T_p be similarity transformations for which

$$(T_p^{-1}A_{l+p}T_p, T_p^{-1}B_{l+p}, CT_p) = (A_{l+p+1}, B_{l+p+1}, C_{l+p+1}),$$

where A_t, B_t, C_t are system-matrices for realizations Σ_t, where $l \leq t \leq k$. We have then that

$$T = T_0T_1T_2T_3 \cdots T_{k-l-3}T_{k-l-2}T_{k-l-1}$$

is a similarity transform satisfying

$$(T^{-1}A_lT, T^{-1}B_l, CT_l) = (A_k, B_k, C_k).$$

Thus, we have that any realization in $\{\Sigma_i\}_{i=n}^{\infty}$ is similar to any other in the sequence. By Proposition 3.2.9 we then have that all realizations share the same properties among $(i), (ii), (iii)$. If Σ_n does not satisfy a property among $(i), (ii), (iii)$, then by the algorithm, Σ_{n+1} would. But since Σ_{n+1} has the same properties as Σ_n, it must be that Σ_n satisfies all three conditions, $(i), (ii)$ and (iii). □

3.2.3 Minimal Dimension of Partial Realizations

Knowing now for certain that the algorithm stops with final state being a realization satisfying $(i) - (iii)$ in Theorem 3.2.1, we can finally embark on proving that minimality is equivalent to satisfying $(i) - (iii)$.

We begin with the following definition.

Definition 3.2.11. The **degree** of a truncated Markov sequence $\mathcal{M} = (M_1, \dots, M_r) \in (\mathbb{K}^{p \times m})^r$ is the smallest possible dimension for a realization of $\mathcal{M}$. It is denoted by $\delta(\mathcal{M})$ or $\delta(M_1, \dots, M_r)$.

Lemma 3.2.12. The degree of a truncated Markov sequence $\mathcal{M} = (M_1, \dots, M_r)$ is bounded below by the quantity

$$\rho(\mathcal{M}) = \rho(M_1, \dots, M_r) \coloneqq \sum_{i+j=r+1} \operatorname{rank} \mathcal{H}_{i,j}(\mathcal{M}) - \sum_{i+j=r} \operatorname{rank} \mathcal{H}_{i,j}(\mathcal{M}).$$

Proof. Let $\Sigma = (A, B, C)$ be an arbitrary realization of $\mathcal{M}$. We then get for $i + j \leq r + 1$ that

$$\mathbf{O}_i(A, C)\mathbf{R}_j(A, B) = \mathcal{H}_{i,j}(\mathcal{M}),$$

where

$$\mathbf{O}_i(A, C) = \begin{pmatrix} C \\ AC \\ \vdots \\ CA^{i-1} \end{pmatrix}, \ \mathbf{R}_j(A, B) = \begin{pmatrix} B & AB & \cdots & A^{j-1}B \end{pmatrix}.$$

We first show that

$$\operatorname{rank} \mathcal{H}_{i,j}(\mathcal{M}) = \dim\{\mathcal{O}_i + \mathcal{R}_j\} - \dim \mathcal{O}_i.$$

This is because

$$\operatorname{rank} \mathcal{H}_{i,j}(\mathcal{M}) = \dim \mathcal{R}_j - \dim\{\mathcal{O}_i \cap \mathcal{R}_j\},$$

by Proposition A.1.9, and that

$$\dim\{\mathcal{O}_i + \mathcal{R}_j\} = \dim \mathcal{O}_i + \dim \mathcal{R}_j - \dim\{\mathcal{O}_i \cap \mathcal{R}_j\}.$$

Thus,

$$\begin{aligned} \operatorname{rank} \mathcal{H}_{i,j}(\mathcal{M}) &= \dim \mathcal{R}_j - \dim\{\mathcal{O}_i \cap \mathcal{R}_j\} \\ &= \dim \mathcal{R}_j + \dim\{\mathcal{O}_i + \mathcal{R}_j\} - \dim \mathcal{O}_i - \dim \mathcal{R}_j \\ &= \dim\{\mathcal{O}_i + \mathcal{R}_j\} - \dim \mathcal{O}_i. \end{aligned}$$

Using this, we get that

$$\begin{aligned} \rho(M) &= \sum_{i+j=r+1} (\dim\{\mathcal{O}_i + \mathcal{R}_j\} - \dim \mathcal{O}_i) - \sum_{i+j=r} (\dim\{\mathcal{O}_i + \mathcal{R}_j\} - \dim \mathcal{O}_i\}) \\ &= \sum_{i=0}^{r-1} (\dim\{\mathcal{O}_{i+1} + \mathcal{R}_{r-i}\}) - \sum_{i=1}^{r-1} (\dim\{\mathcal{O}_i + \mathcal{R}_{r-i}\}) - \dim \mathcal{O}_r \\ &= \dim\{\mathcal{O}_1 + \mathcal{R}_r\} - \dim \mathcal{O}_r + \sum_{i=1}^{r-1} (\dim\{\mathcal{O}_{i+1} + \mathcal{R}_{r-i}\}) - \sum_{i=1}^{r-1} (\dim\{\mathcal{O}_i + \mathcal{R}_{r-i}\}) \\ &= \dim\{\mathcal{O}_1 + \mathcal{R}_r\} - \dim \mathcal{O}_r - \sum_{i=1}^{r-1} (\dim\{\mathcal{O}_i + \mathcal{R}_{r-i}\} - \dim\{\mathcal{O}_{i+1} + \mathcal{R}_{r-i}\}). \end{aligned}$$

Note that

$$\sum_{i=1}^{r-1} \left(\dim\{\mathcal{O}_i + \mathcal{R}_{r-i}\} - \dim\{\mathcal{O}_{i+1} + \mathcal{R}_{r-i}\}\right) \geq 0$$

since $\mathcal{O}_j \supseteq \mathcal{O}_{j+1}$. Thus, it is clear that

$$\rho(\mathcal{M}) \leq \dim \Sigma,$$

since $\mathcal{O}_1 + \mathcal{R}_r \subseteq \mathbb{K}^{\dim \Sigma}$ has dimension at most $\dim \Sigma$. Taking the minimum dimension among realizations of $\mathcal{M}$, we finally get that

$$\rho(\mathcal{M}) \leq \delta(\mathcal{M}).$$

□

We now have everything need in order to prove Theorem 3.2.1.

Proof of Theorem 3.2.1. Suppose $\Sigma = (A, B, C)$ is a realization of $\mathcal{M}$ satisfying $(i) - (iii)$. We calculated in the proof above that

$$\rho(\mathcal{M}) = \dim\{\mathcal{O}_1 + \mathcal{R}_r\} - \dim \mathcal{O}_r - \sum_{i=1}^{r-1} \left(\dim\{\mathcal{O}_i + \mathcal{R}_{r-i}\} - \dim\{\mathcal{O}_{i+1} + \mathcal{R}_{r-i}\}\right).$$

Since Σ satisfies $(i) - (iii)$, we have that

$$\begin{aligned} &\mathcal{R}_r = \mathbb{K}^{\dim \Sigma}, \\ &\mathcal{O}_r = \{\vec{0}\}, \\ &\mathcal{O}_j \subseteq \mathcal{R}_{r-j}, \quad j = 1, \ldots, r-1, \end{aligned}$$

so,

$$\rho(\mathcal{M}) = \dim \Sigma - 0 - \sum_{i=1}^{r-1} \left(\dim \mathcal{R}_{r-i} - \dim \mathcal{R}_{r-i}\right) = \dim \Sigma,$$

since $\mathcal{O}_i \supseteq \mathcal{O}_{i+1}$ for $i = 1, \ldots, r-1$ in particular. Thus, we have that $\delta(\mathcal{M}) \geq \dim \Sigma$, and by minimality, we must have that $\dim \Sigma = \delta(\mathcal{M})$.
Suppose now that $\Sigma = (A, B, C)$ is a minimal realization of $\mathcal{M}$. By compressing via Lemmas 3.2.4, 3.2.5, 3.2.3, we can achieve a realization Σ' satisfying $(i) - (iii)$. By minimality, we must have that $\dim \Sigma' = \dim \Sigma$, meaning that triples are similar. Thus, Σ satisfies $(i) - (iii)$. □

Example 13 (Minimal Realization of Truncated Fibonacci)**.** Consider the scalar sequence $\mathcal{F} = (1, 1, 2)^\top$. A realization for this sequence is $\Sigma = (A, B, C)$ where

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}.$$

This is just the observability form realization of the Markov sequence

$$M_i = \begin{cases} 1 & \text{if } i = 1, \\ 1 & \text{if } i = 2, \\ 2 & \text{if } i = 3, \\ 0 & \text{otherwise.} \end{cases}$$

We have that

$$\begin{aligned}\delta(\mathcal{F}) &= \sum_{i+j=4} \operatorname{rank} \mathcal{H}_{i,j}(\mathcal{F}) - \sum_{i+j=3} \operatorname{rank} \mathcal{H}_{i,j}(\mathcal{F}) \\ &= \operatorname{rank} \mathcal{H}_{1,3}(\mathcal{F}) + \mathcal{H}_{3,1}(\mathcal{F}) + \operatorname{rank} \mathcal{H}_{2,2}(\mathcal{F}) - \operatorname{rank} \mathcal{H}_{1,2}(\mathcal{F}) - \operatorname{rank} \mathcal{H}_{2,1}(\mathcal{F}),\end{aligned}$$

where

$$\begin{aligned}\mathcal{H}_{1,3}(\mathcal{F}) &= \begin{pmatrix} 1 & 1 & 2 \end{pmatrix}, \\ \mathcal{H}_{3,1}(\mathcal{F}) &= \mathcal{H}_{1,3}(\mathcal{F})^\top, \\ \mathcal{H}_{2,2}(\mathcal{F}) &= \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \\ \mathcal{H}_{1,2}(\mathcal{F}) &= \begin{pmatrix} 1 & 1 \end{pmatrix} \\ \mathcal{H}_{2,1}(\mathcal{F}) &= \mathcal{H}_{1,2}(\mathcal{F})^\top,\end{aligned}$$

since $\mathcal{F}$ is a scalar sequence. Thus, we have that

$$\delta(\mathcal{F}) = 1 + 1 + 2 - 1 - 1 = 2.$$

Since Σ has dimension 3, it is not minimal, so we check which properties from Theorem 3.2.1 fail. We have that

$$\begin{aligned}\mathbf{R}_3(A, B) &= \begin{pmatrix} B & AB & A^2B \end{pmatrix} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 0 \end{pmatrix}, \\ \mathbf{O}_3(A, C) &= \begin{pmatrix} C \\ CA \\ CA^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3,\end{aligned}$$

which means that $\mathcal{O}_3 = \{\vec{0}\}$ and $\mathcal{R}_3 = \mathbb{R}^3$. We also have that

$$\begin{aligned}\ker C &= \operatorname{span} \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}, \\ \ker \begin{pmatrix} C \\ CA \end{pmatrix} &= \operatorname{span} \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}, \\ \operatorname{Im}(B) &= \operatorname{span} \left\{ \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right\}, \\ \operatorname{Im} \begin{pmatrix} B & AB \end{pmatrix} &= \operatorname{span} \left\{ \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \right\}.\end{aligned}$$

Therefore, we see that property (iii) does not hold for $j = 1$ nor $j = 2$. We perform a decomposition in accordance with Lemma 3.2.5. We fix $j = 1$ and set

$$\begin{aligned}\mathcal{X}_1 &= \ker \begin{pmatrix} C \\ CA \end{pmatrix}, \\ \mathcal{Y}_0 &= \operatorname{Im} \begin{pmatrix} B & AB \end{pmatrix}.\end{aligned}$$

Then, we have that

$$T = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 0 \end{pmatrix},$$

and that

$$\begin{aligned}F = T^{-1}AT &= \begin{pmatrix} 2 & 0 & -8 \\ -1 & 0 & 4 \\ 1 & 1 & -2 \end{pmatrix}, \\ G = T^{-1}B &= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \\ H = CT &= \begin{pmatrix} 0 & 1 & 1 \end{pmatrix}.\end{aligned}$$

Note that Lemma 3.2.5 says the dimension of the new realization should be $3 - \dim \mathcal{X}_1 = 3 - 1 = 2$, so we have that

$$F_{00} = \begin{pmatrix} 0 & 4 \\ 1 & -2 \end{pmatrix}, G_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, H_0 = \begin{pmatrix} 1 & 1 \end{pmatrix}.$$

Simple calculation shows that the triple $\Sigma_0 = (F_{00}, G_0, H_0)$ indeed is a realization of $\mathcal{F}$ and since its dimension is equal to the degree of $\mathcal{F}$, we are done.

3.2.4 Uniqueness of Minimal Partial Realizations

We have previously seen that the solution to the minimal partial realization is not necessarily unique. A sufficient and necessary condition for uniqueness (up to similarity) is given by the following.

Theorem 3.2.13. *Let $\mathcal{M} = (M_1, \ldots, M_r) \in (\mathbb{K}^{p\times m})^r$ be a truncated Markov sequence. The quantities*

$$\begin{aligned}\alpha(\mathcal{M}) = \alpha(M_1, \ldots, M_r) &\coloneqq \min\{i : \operatorname{rank} \mathcal{H}_{i,r-i}(\mathcal{M}) = \operatorname{rank} \mathcal{H}_{i+1,r-i}(\mathcal{M}), 1 \le i < r\}, \\ \beta(\mathcal{M}) = \beta(M_1, \ldots, M_r) &\coloneqq \min\{i : \operatorname{rank} \mathcal{H}_{r-i,i}(\mathcal{M}) = \operatorname{rank} \mathcal{H}_{r-i,i+1}(\mathcal{M}), 1 \le i < r\},\end{aligned}$$

where we take $\min \emptyset$ to be r, satisfy the inequality

$$\alpha(\mathcal{M}) + \beta(\mathcal{M}) \le r$$

if and only if all minimal realizations of $\mathcal{M}$ are similar.

Notation 3.2.14. When it is not ambiguous, we will use the notation α and β instead of $\alpha(\mathcal{M})$ and $\beta(\mathcal{M})$.

The complete proof of the statement above is only cited to in [GKL87] and involves a lot of interesting algebra and topology that is beyond the scope of this thesis. The interested reader can find the proof of the above statement (formulated in a *very* different way) in [Bos83] and its relevant sources.

In this thesis, we provide an alternative proof for the sufficiency-part of the above theorem, utilizing the following lemma.

Lemma 3.2.15. Let $\mathcal{M} = (M_1, \ldots, M_r) \in (\mathbb{K}^{p\times n})^r$ be a truncated Markov sequence satisfying $\alpha(\mathcal{M}) + \beta(\mathcal{M}) \leq r + 1$. Then

1. $\delta \coloneqq \delta(\mathcal{M}) = \operatorname{rank} \mathcal{H}_{\alpha,\beta}(\mathcal{M})$, and
2. (A, B, C) is a minimal realization of $\mathcal{M}$ if and only if

$$\ker \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{\alpha-1} \end{pmatrix} = \{\vec{0}\},$$

$$\operatorname{Im} \begin{pmatrix} B & AB & \cdots & A^{\beta-1}B \end{pmatrix} = \mathbb{K}^{\delta}.$$

Proof. [GKL87]. □

Proof of Theorem 3.2.13 (Sufficiency). We assume that $\alpha + \beta \leq r$ and let $\Sigma_1 = (A, B, C)$ be a minimal realization of $\mathcal{M}$. We denote the Markov sequence of W_{Σ_1} by $\mathcal{M}' = \{\mathcal{M}'_i\}_{i=1}^{\infty}$. We have by Lemma 2.2.4 that

$$\mathcal{H}_{\alpha,\beta}(\mathcal{M}) = \mathcal{H}_{\alpha,\beta}(\mathcal{M}') = \mathbf{O}_\alpha(A, C)\mathbf{R}_\beta(A, B).$$

By Lemma 3.2.15, we have that $\operatorname{rank} \mathcal{H}_{\alpha,\beta}(\mathcal{M}) = \delta$, and since (A, B, C) is minimal, we also have by 2.2.4 that

$$\operatorname{rank} \mathbf{O}_\alpha(A, C) = \delta \text{ and } \operatorname{rank} \mathbf{R}_\beta(A, B) = \delta.$$

Thus, $\mathcal{H}_{\alpha,\beta}(\mathcal{M}) = \mathbf{O}_\alpha(A, C)\mathbf{R}_\beta(A, B)$, and thus is a full-rank factorization of $\mathcal{H}_{\alpha,\beta}(\mathcal{M})$. Consider now another minimal realization $\Sigma_2 = (F, G, H)$ of $\mathcal{M}$ and the Markov sequence of W_{Σ_2}, which we denote by $\mathcal{M}'' = \{\mathcal{M}''_i\}_{i=1}^{\infty}$. This triple yields another full-rank factorization of $\mathcal{H}_{\alpha,\beta}(\mathcal{M}) = \mathcal{H}_{\alpha,\beta}(\mathcal{M}'')$, namely

$$\mathcal{H}_{\alpha,\beta}(\mathcal{M}) = \mathbf{O}_\alpha(F, H)\mathbf{R}_\beta(F, G).$$

By Proposition A.1.4, there is a $T \in \mathrm{GL}_\delta(\mathbb{K})$ such that

$$\mathbf{O}_\alpha(A, B) = \mathbf{O}_\alpha(F, H)T \text{ and } \mathbf{R}_\beta(A, B) = T^{-1}\mathbf{R}_\beta(F, G).$$

By looking at the equalities block-entry-wise, we find that $B = T^{-1}G$ and $C = HT$. Recall that

$$\mathcal{H}^1_{\alpha+1,\beta}(\mathcal{M}) = \mathcal{H}^1_{\alpha+1,\beta}(\mathcal{M}') = \mathcal{H}^1_{\alpha+1,\beta}(\mathcal{M}'')$$

is well-defined since the last block-entry $M_{\alpha+\beta}$ exists and that

$$\mathcal{H}^1_{\alpha+1,\beta}(\mathcal{M}) = \mathbf{O}_\alpha(A, C)A\mathbf{R}_\beta(A, B) = \mathbf{O}_\alpha(F, H)F\mathbf{R}_\beta(F, G).$$

Thus, we have, since left-and-right inverses exist by the matrices being full-rank, that

$$A = \mathbf{O}_\alpha(A, C)^{-L}\mathbf{O}_\alpha(F, H)F\mathbf{R}_\beta(F, G)\mathbf{R}_\beta(A, B)^{-R}.$$

Since $T^{-1} = \mathbf{O}_\alpha(A, C)^{-L}\mathbf{O}_\alpha(F, H)$ and $T = \mathbf{R}_\beta(F, G)\mathbf{R}_\beta(A, B)^{-R}$ from the above relations, we have that

$$(A, B, C) = (T^{-1}FT, T^{-1}G, HT).$$

Since (A, B, C) and (F, G, H) were chosen arbitrarily, we have that similarity holds for all minimal realizations of $\mathcal{M}$. □

Remark 3.2.16. Note that by proving above that Σ_1 and Σ_2 are similar, we have proven that they in fact generate the exact same Markov sequence. In other words, that $W_{\Sigma_1} = W_{\Sigma_2}$. This is significant since we apriori only knew that they matched up to and including the rth Markov parameter, i.e. the last parameter in $\mathcal{M}$. Another interpretation is that, in the case of $\alpha + \beta \leq r$, there is a unique Markov sequence of rank $\delta(\mathcal{M})$ that extends $\mathcal{M}$. This connects to Proposition 2.3.2, which is a similar result.

Chapter 4

Approximate Realization and Structured Total Least Squares

The approximate realization problem essentially asks for a transfer function whose first r Markov parameters approximately match up with a truncated Markov sequence, as opposed to exactly matching up with it. In this thesis, we study this problem in the real SISO-case, i.e., the case where $p = m = 1$ and $\mathbb{K} = \mathbb{R}$.

4.1 Approximate Realization

We now present the problem a bit formally.

Problem 3 (Approximate Realization)**.** Given a truncated Markov sequence $\mathcal{M} = (M_1, M_2, \ldots, M_r)^\top$ of real numbers, find $n \in \mathbb{Z}^+$ and matrices $A \in \mathbb{R}^{n\times n}, b \in \mathbb{R}^{n\times 1}, c \in \mathbb{R}^{1\times n}$ such that

1. $\sum_{i=1}^{r} |M_i - cA^{i-1}b|^2$ is small,
2. for $\Sigma = (A, b, c)$, we have $\dim \Sigma < \delta(\mathcal{M})$.

Remark 4.1.1. The reason for not specifying how small $\sum_{i=1}^{r} |M_i - cA^{i-1}b|^2$ should be by setting an upper bound is that it is not all that clear if for every error tolerance $\epsilon > 0$ there is a triple (A, b, c) such that

$$\sum_{i=1}^{r} |M_i - cA^{i-1}b|^2 < \epsilon.$$

Investigation into which error tolerances admit triples (A, b, c) and which do not is beyond the scope of this thesis. Furthermore, nothing is said about $\sum_{i=1}^{r} |M_i - cA^{i-1}b|^2$ being minimal among systems with dimension lower than $\delta(\mathcal{M})$. This is choice is strategic since it otherwise would be a problem of rank-constrained minimization, which is also beyond the scope of this thesis. Nothing is said concerning how much lower the dimension of an approximate realization should be for the same reason.

We say that a triple $\hat{\Sigma} = (A, b, c)$ is an **approximate realization** for $\mathcal{M} = (M_1, \ldots, M_r)^\top$ if it solves Problem 3 for the same truncated Markov sequence. All other notions and concepts concerning partial realizations stay the same.

4.1.1 Kung's Algorithm

One method of solving the approximate realization problem for instances with an odd number of Markov parameters is via *Kung's algorithm*. The algorithm has two main steps. The first step involves a *low-rank approximation* of a Hankel matrix made up from the Markov parameters. A low rank approximation of a matrix M is a matrix $\hat{M}$ which is of lower rank than M and is close in some norm sense to M, i.e., $\|M - \hat{M}\|$ is small. One way of obtaining a low rank approximation in the Frobenius norm is via the following lemma which details how to construct a best possible approximation with respect to some upper bound on the rank.

Lemma 4.1.2. Let $D = U\Sigma V^\top$, be the singular value decomposition of D and suppose $U, \Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_p), V$ are partitioned as

$$U = \begin{pmatrix} U_1 & U_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_1 & \mathbf{0} \\ \mathbf{0} & \Sigma_2 \end{pmatrix}, \quad V = \begin{pmatrix} V_1 & V_2 \end{pmatrix}$$

where U_1 and V_1 are the first m columns of U and V respectively and $\Sigma_1 = \operatorname{diag}(\sigma_1, \ldots, \sigma_m)$. Then

$$D^* = U_1 \Sigma_1 V_1^\top$$

solves the program to

$$\text{minimize over } \hat{D} \quad \|D - \hat{D}\|_F \quad \text{subject to} \quad \operatorname{rank} \hat{D} \leq m.$$

Proof. [Mar11, Appendix B] □

The second step of the algorithm is the actual construction of the system-matrices of the approximate realization. It resembles the construction of the system matrices in the proof of Theorem 2.2.6 in the way that we factor the low-rank approximation of the Hankel matrix into a full-rank factorization. A huge detail to note is that the low-rank approximation need not be Hankel, so we lose some structure in the sense that our factors may not be observability-and-controllability matrices to any system, and that the constructed realization may or may not reproduce the low-rank approximation as its Hankel matrix. Furthermore, knowing how good/ bad the approximate realization is parameter-wise is not possible since, again, the low-rank approximation may not be a Hankel matrix, meaning there is no truncated Markov sequence that produces it.

We now present the construction algorithm in the form of an example. The algorithm is presented as in [PCM86].

Example 14. Let $\mathcal{M} = (M_1, \ldots, M_{2r-1})$ be a truncated Markov sequence. We form the singular value decomposition of its (r, r)th Hankel matrix

$$\mathcal{H}_{r,r}(\mathcal{M}) = U\Sigma V^\top,$$

and then form the low-rank approximation

$$H = U_1 \Sigma_1 V_1^\top$$

according to Lemma 4.1.2, for m smaller than $\delta(\mathcal{M})$. Then, note for $O = U\Sigma^{1/2}$ and $R = \Sigma^{1/2} V^\top$, with $\Sigma^{1/2} = \operatorname{diag}(\sqrt{\sigma_1}, \ldots, \sqrt{\sigma_m})$, that

$$H = OR$$

is a full-rank factorization of H. We let B be the first column of R and C be the first row of O. The matrix A is retrieved through the following reasoning. If (F, G, H) is an $n-$dimensional realization of $\mathcal{M}$, then

$$F\mathbf{R}(F,G)_f = \mathbf{R}(F,G)_l, \quad \mathbf{O}(F,H)_f F = \mathbf{O}(F,G)_l$$

where subscript f means first $n-1$ rows or columns and l means the last $n-1$ rows or columns. In our case, O and R play the roles of $\mathbf{R}(F,G)$ and $\mathbf{O}(F,H)$, so we can find A by solving

$$AR_f = R_l \quad \text{or} \quad O_f A = O_l.$$

Note that O_f is full-rank since it the first $n-1$ vectors of $O = U\Sigma^{1/2}$, whose columns are all linearly independent. Thus, we can do the following

$$O_f^\top O_f A = O_f^\top O_l \implies A = (O_f^\top O_f)^{-1} O_f^\top O_l,$$

by Proposition A.1.3. Thus, $A = (O_f^\top O_f)^{-1} O_f^\top O_l$ is a suitable matrix for our approximate realization (A, B, C).

4.1.2 Towards a Structure-preserving Method

What we would like to keep in the low-rank approximation is the Hankel structure of the original matrix. Such an approximation is called a Hankel low-rank approximation. In the context of our problem, we want to find an approximation of the (r, r)th Hankel matrix of a real scalar sequence $\mathcal{M} = (M_1, \ldots, M_{2r-1})^\top$ that is of lower rank and is Hankel. The method we intend to use assumes that the original Hankel matrix is non-singular. This means that the method we are about to discuss is not a universal way to find approximate realizations. We therefore begin by first identifying a class of truncated Markov sequences we can apply the method on.

Theorem 4.1.3. *Let $r \in \mathbb{Z}^+$. For a truncated sequence of scalar Markov parameters $\mathcal{M} = (M_1, \ldots, M_{2r-1})^\top \in \mathbb{R}^{2r-1}$ we have that*

$$\det \mathcal{H}_{r,r}(\mathcal{M}) \neq 0 \iff \delta(\mathcal{M}) = r.$$

Proof. Suppose $\det \mathcal{H}_{r,r}(\mathcal{M}) \neq 0$. This means that $\operatorname{rank} \mathcal{H}_{r,r}(\mathcal{M}) = r$. If $\Sigma = (A, b, c)$ is a minimal realization of $\mathcal{M}$ with dimension less than r we would have, by Lemma 2.2.5, for $s, t \geq \dim \Sigma$ that

$$\operatorname{rank} \mathcal{H}_{s,t}(\mathcal{M}') = \dim \Sigma < r$$

where $\mathcal{M}'$ is the infinite extension of $\mathcal{M}$ given by

$$\mathcal{M}'_i = cA^{i-1}b, \quad i \in \mathbb{Z}^+.$$

In particular, since $r \geq \dim \Sigma$, we have that

$$\operatorname{rank} \mathcal{H}_{r,r}(\mathcal{M}') < r.$$

This is contradictory since $\mathcal{H}_{r,r}(\mathcal{M}') = \mathcal{H}_{r,r}(\mathcal{M})$.
Suppose now that $\delta(\mathcal{M}) = r$. Then there is an $r-$dimensional realization $\Sigma = (A, b, c)$ of $\mathcal{M}$. Thus, we have, again by Lemma 2.2.5, for $s, t \geq r$ that

$$\operatorname{rank} \mathcal{H}_{s,t}(\mathcal{M}') = r$$

where $\mathcal{M}'$ is the infinite extension of $\mathcal{M}$ given by

$$\mathcal{M}'_i = cA^{i-1}b, \quad i \in \mathbb{Z}^+.$$

In particular, for $s = t = r$ we have that

$$\operatorname{rank} \mathcal{H}_{r,r}(\mathcal{M}') = r.$$

Since $\mathcal{H}_{r,r}(\mathcal{M}') = \mathcal{H}_{r,r}(\mathcal{M})$, we get that $\mathcal{H}_{r,r}(\mathcal{M})$ is non-singular. □

With the above result, we see that low-rank approximation via the method to be presented is possible for the (r, r)th Hankel matrix of truncated Markov sequences of length $2r - 1$ that have degree equaling to r. We detail now through two examples how this can be used in combination with some of the discussed methods of partial realization to yield approximate realizations.

Example 15. Let $\mathcal{M} = (M_1, \dots, M_{2r-1})^\top$ be a truncated real Markov sequence with $\delta(\mathcal{M}) = r$. Let now H be a Hankel low-rank approximation of $\mathcal{H}_{r,r}(\mathcal{M})$ in the Frobenius norm, i.e.

- $H \in \mathbb{R}^{r\times r}$ is Hankel,
- $\operatorname{rank} H < \operatorname{rank} \mathcal{H}_{r,r}(\mathcal{M})$
- $\|H - \mathcal{H}_{r,r}(\mathcal{M})\|_F$ is small.

The entries of H constitute an approximation of $\mathcal{M}$, which we will denote $\mathcal{N} = (N_1, N_2, \dots, N_{2r-1})^\top$, so $H = \mathcal{H}_{r,r}(\mathcal{N})$. We have that $\delta(\mathcal{N}) < r$ since we can apply the methods of residue interpolation to find a realization of dimension $\operatorname{rank} H < r$. Thus, by finding a low-rank approximation of $\mathcal{H}_{r,r}(\mathcal{M})$, we get a sequence close to $\mathcal{M}$ with smaller degree, meaning we can realize the new sequence and by the Compression Algorithm get an approximate realization of $\mathcal{M}$.

Example 16. Let $\mathcal{M} = (M_1, \dots, M_{2r-1})^\top$ again be a finite real sequence with $\delta(\mathcal{M}) = r$. We can find a minimal realization of $\mathcal{M}$ by finding a full-rank decomposition of $\mathcal{H}_{r,r}(\mathcal{M})$ and using the methods from rational interpolation. This (partial) realization will be canonical, but aslo minimal since $\delta(M) = \operatorname{rank} \mathcal{H}_{r,r}(\mathcal{M})$. If we instead had a Hankel low-rank approximation H of $\mathcal{H}_{r,r}(\mathcal{M})$, we could instead factor H to a full-rank decomposition and then use the same methods to yield a realization of the finite Markov sequence H is made up of. This realization would be of lower degree than that of $\mathcal{H}_{r,r}(\mathcal{M})$ and the difference in Markov parameters would be small. Thus, we get an approximate realization of $\mathcal{M}$.

4.2 Hankel Total Least Squares

As was hinted at in Example 15, the Hankel low-rank approximation problem for a full-rank matrix H can be posed as one to

$$\text{minimize over } \hat{H} \quad \|H - \hat{H}\|_F \quad \text{subject to} \quad \begin{cases} \hat{H} \text{ Hankel}, \\ \operatorname{rank} \hat{H} < \operatorname{rank} H. \end{cases} \tag{HLRA}$$

The rank constraint in this formulation is difficult to deal with, so we will not be solving this particular problem in this thesis. We instead formulate a similar problem without a rank constraint, by treating it as a problem of **structured total least squares** (STLS). To do so, we need do define what is meant by "structure" in the structured total least-square formulation.

Definition 4.2.1. A map $\mathcal{B}\colon \mathbb{R}^n \to \mathbb{R}^{p\times m}$ is called an **affine matrix-function** if there are matrices $B_0, B_1, \dots, B_n \in \mathbb{K}^{p\times m}$ such that

$$\mathcal{B}(x) = B_0 + \sum_{i=1}^{n} x_i B_i \quad \text{for all } x \in \mathbb{R}^n,$$

where $x = (x_1, x_2, \dots, x_n)^\top$.

Example 17. Let $h = (h_1, h_2, \dots, h_{2n-1})^\top \in \mathbb{R}^{2n-1}$ be an arbitrary vector. The (n,n)th Hankel matrix of h, $\mathcal{H}_{n,n}(h)$, can be seen as the application of an affine matrix-function, namely the (n,n)th Hankel affine matrix-function $\mathcal{H}_{n,n} : \mathbb{R}^{2n-1} \to \mathbb{R}^{n\times n}$. We have that

$$\mathcal{H}_{n,n}(h) = \sum_{i=1}^{2n-1} h_i H_i,$$

where $H_i \in \mathbb{R}^{n\times n}$, $i \in \{1, \dots, 2n-1\}$ are matrices defined entry-wise by

$$(H_i)_{k,l} = \begin{cases} 1 & \text{if } k + l = i + 1, \\ 0 & \text{otherwise.} \end{cases}$$

As an illustration, the first three matrices are

$$H_1 = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \; H_2 = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \; H_3 = \begin{pmatrix} 0 & 0 & 1 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

By checking for the properties, one will find that the Hankel affine matrix-function is linear, one-to-one and maps onto the subspace of Hankel matrices in $\mathbb{R}^{n\times n}$.

Notation 4.2.2. To reduce clutter, we will from here on out denote the (n,n)th Hankel affine matrix-function by $\mathcal{H}_n$.

With the concept of a Hankel affine matrix-function and noting that rank deficiency for square matrices is equivalent to having a non-trivial null-space, we can formulate for a full-rank Hankel matrix $\mathcal{H}_n(h)$ the **Hankel total least squares** problem.

Problem 4 (Hankel Total Least Squares)**.** Let $h \in \mathbb{R}^{2r-1}$ be such that $\mathcal{H}_r(h)$ is non-singular. Find $((\hat{h}^*)^\top, (y^*)^\top)^\top \in \mathbb{R}^{3r-1}$ that solves the program to

$$\text{minimize over } \hat{h} \text{ and } y \quad \sum_{i=1}^{2r-1} (h_i - \hat{h}_i)^2 \quad \text{subject to} \quad \begin{cases} \hat{h} \in \mathbb{R}^{2r-1}, \\ y \in \mathbb{R}^r, \\ \mathcal{H}_r(\hat{h})y = \vec{0}, \\ y^\top y = 1. \end{cases} \tag{HTLS}$$

Remark 4.2.3. Replacing the Hankel affine matrix-function with a general affine matrix-function yield the general STLS. As in the specific case, we always assume that $\mathcal{B}(h)$ is full-rank, so that h is not in the feasible set.

Solutions, when they exist, to HTLS may not solve the Hankel low-rank approximation problem as formulated as HLRA, but make for good approximations nonetheless. The advantage of this formulation though, is that it is an equality-constrained optimization problem with objective function and constraint function both being continuously differentiable, which means we can use the method of Lagrange multipliers provided the LICQ (see Definition A.2.1) holds at all relevant points.

The method of multipliers is the main tool used in [Moo94] to prove that another set of equations can be used to find global solutions to a general STLS. In this thesis, we will focus on the case where the STLS is the HTLS, since further analysis shows that it has solutions and that the LICQ is satisfied at the relevant points.

4.2.1 Existence of Minimizers

Notation 4.2.4. We will, for the rest of the thesis, identify points in $\mathbb{R}^{n+m}$ by points in $\mathbb{R}^n \times \mathbb{R}^m$ and write vectors $x \in \mathbb{R}^{n+m}$ as (x_1, x_2), where $x_1 \in \mathbb{R}^n$ and $x_2 \in \mathbb{R}^m$. This is purely notational so when we deal with functions $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, we are actually talking about a function $f : \mathbb{R}^{n+m} \to \mathbb{R}$, and when we for $x_1 \in \mathbb{R}^n$, $x_2 \in \mathbb{R}^m$ write $(x_1, x_2) \in \mathbb{R}^{n+m}$, we actually mean that $(x_1^\top, x_2^\top)^\top \in \mathbb{R}^{n+m}$.

It is not all that clear that HTLS has solutions. One might naively want to argue by the *Extreme Value Theorem* since the objective function $f(b, y) = \|a - b\|_2^2$ is continuous, but the problem is that the feasible set

$$\mathcal{F} = \{(b, y) \in \mathbb{R}^{2r-1} \times \mathbb{R}^r : \mathcal{H}_r(b)y = \vec{0},\ y^\top y = 1\}$$

is not bounded for $r \geq 2$. To show this, note first that it is non-empty, since $(b, e_1) \in \mathcal{F}$, where

$$\begin{aligned} e_1 &= (1, 0, 0, \ldots, 0)^\top \in \mathbb{R}^r \\ b &= (\underbrace{0, 0, \ldots, 0}_{r \text{ entries}}, 1, 1, 1, \ldots, 1)^\top \in \mathbb{R}^{2r-1}. \end{aligned}$$

By linearity of $\mathcal{H}$, we have for each $\lambda \in \mathbb{R}$ that $(b', e_1) \in \mathcal{F}$, where

$$b' = \lambda \frac{b}{\|b\|_2}.$$

For the case when $r = 1$ we have that $r = 2r - 1$. The HTLS is not meaningful since the Hankel map is the identity, meaning the feasible set is

$$\mathcal{F} = \{(b, y) \in \mathbb{R} \times \mathbb{R} : \mathcal{H}_1(b) = by = 0,\ y^2 = 1\} = \{(0, 1), (0, -1)\}.$$

Thus $\hat{h} = 0$ is the only approximation for $h \neq 0$ which is of lower dimension. In the context of approximate realization, we have a degenerate case for $r = 1$, since if we want to find approximate realizations for only one Markov parameter, it would have to be zero-dimensional.

We now introduce concepts and a theorem from [Gra17] that can be used to prove existence of solutions for HTLS. Instead of looking at compactness of the domain, we leverage nice properties of the objective function. We need the following two definitions.

Definition 4.2.5. A function $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ is **lower semi-continuous** if for every $x \in D$ and every sequence $x_k \in D$ converging to x, we have that

$$f(x) \leq \liminf_{k \to \infty} f(x_k)$$

Example 18. Continuous functions are lower semi-continuous.

Definition 4.2.6. A function $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ is **coercive** if for every sequence $x_k \in D$ such that $\|x_k\|_2 \to \infty$, we have that $f(x_k) \to \infty$.

Example 19. The objective function $f(b, y) = \|a - b\|_2^2$ is coercive.

We now state and prove an existence result, and then prove that the HTLS has at least one solution.

Theorem 4.2.7. *Suppose $\mathcal{F}$ is a non-empty closed subset of $D \subseteq \mathbb{R}^n$ and that $f : D \to \mathbb{R} \cup \{+\infty\}$ is lower semi-continuous and coercive. The program to*

$$\textit{minimize over } x \quad f(x) \quad \textit{subject to} \quad x \in \mathcal{F},$$

admits a solution.

Proof. If $f(x) = +\infty$ for all $x \in \mathcal{F}$, we have nothing to do, so assume that there is one $x' \in \mathcal{F}$ such that $f(x') < +\infty$. Thus, we have that $f^* := \inf_{x \in \mathcal{F}} f(x) \leq f(x') < +\infty$. Let $\{x_k\}_k$ be a sequence of points for which the images converge to f^*. What this means is that $f(x_k) \not\to \infty$, meaning the sequence has to be bounded, by coercivity of f. By the *Bolzano-Weierstrass Theorem*, there is a subsequence $\{x_{k_m}\}_m$ that converges to some $x^* \in \mathbb{R}^n$. Since $\mathcal{F}$ is closed, we have $x^* \in \mathcal{F}$. Thus,

$$f^* = \lim_{k \to \infty} f(x_k) = \lim_{k_m \to \infty} f(x_{k_m}) = \liminf_{k_m \to \infty} f(x_{k_m}) \geq f(x^*),$$

meaning $f(x^*) = f^*$. Hence, f attains a minimum in $\mathcal{F}$. □

Corollary 8. The HTLS has a solution.

Proof. The HTLS program can be written as a program to

$$\text{minimize over } \hat{h} \text{ and } y \quad f(\hat{h}, y) \quad \text{subject to} \quad g(\hat{h}, y) = \vec{0},$$

where $f : \mathbb{R}^{2r-1} \times \mathbb{R}^r \to \mathbb{R}$ and $g : \mathbb{R}^{2r-1} \times \mathbb{R}^r \to \mathbb{R}^{r+1}$ are given by

$$f(\hat{h}, y) = \|h - \hat{h}\|_2^2, \quad g(\hat{h}, y) = \begin{pmatrix} \mathcal{H}_r(\hat{h})y \\ y^\top y - 1 \end{pmatrix}.$$

Thus, it is clear that our feasible set is $\mathcal{F} = \ker g$. By later analysis of g, we will see that g has continuous partials and thus is differentiable. Hence, it is continuous, meaning that $\mathcal{F}$ is closed.

Our objective function f is continuous and also coercive. Thus, the statement follows from the previous theorem. □

4.2.2 Existence of Multipliers

Since we are dealing with an equality-constrained problem, we would like to use the Lagrange multiplier method. To do this, we need to prove that multipliers exist by showing that the LICQ is satisfied.

Lemma 4.2.8. The LICQ for HTLS is satisfied at every point in the feasible set. In particular, it is satisfied at every local minimum.

Proof. We again use the compact form of writing the problem, which is to

$$\text{minimize over } \hat{h} \text{ and } y \quad f(\hat{h}, y) \quad \text{subject to} \quad g(\hat{h}, y) = \vec{0}.$$

Our goal here is to show that $\operatorname{rank} \nabla g(\eta, \upsilon) = r + 1$ for an arbitrary feasible point (η, υ). The gradient of g at (η, υ) is given by

$$\nabla g(\eta, \upsilon) = \left(\begin{array}{cccc|c} H_1\upsilon & H_2\upsilon & \cdots & H_{2r-1}\upsilon & \mathcal{H}_r(\eta) \\ \hline 0 & 0 & \cdots & 0 & 2\upsilon^\top \end{array}\right) \in \mathbb{R}^{(r+1)\times(3r-1)}.$$

A sub-matrix of particular interest is

$$S = \begin{pmatrix} H_1\upsilon & H_2\upsilon & \cdots & H_{2r-1}\upsilon \end{pmatrix} \in \mathbb{R}^{r\times(2r-1)}.$$

We have that the first r columns are given by

$$H_1\upsilon = \begin{pmatrix} \upsilon_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \; H_2\upsilon = \begin{pmatrix} \upsilon_2 \\ \upsilon_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \; H_3\upsilon = \begin{pmatrix} \upsilon_3 \\ \upsilon_2 \\ \upsilon_1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, H_r\upsilon = \begin{pmatrix} \upsilon_r \\ \upsilon_{r-1} \\ \upsilon_{r-2} \\ \vdots \\ \upsilon_1 \end{pmatrix}$$

and the rest of the columns are

$$H_{r+1}\upsilon = \begin{pmatrix} 0 \\ \upsilon_r \\ \upsilon_{r-1} \\ \vdots \\ \upsilon_2 \end{pmatrix}, \; H_{r+2}\upsilon = \begin{pmatrix} 0 \\ 0 \\ \upsilon_r \\ \vdots \\ \upsilon_3 \end{pmatrix}, \ldots, H_{2r-1}\upsilon = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ \upsilon_r \end{pmatrix}.$$

In other words, we obtain

$$S = \begin{pmatrix} \upsilon_1 & \upsilon_2 & \upsilon_3 & \cdots & \upsilon_r & 0 & 0 & \cdots & 0 \\ 0 & \upsilon_1 & \upsilon_2 & \cdots & \upsilon_{r-1} & \upsilon_r & 0 & \cdots & 0 \\ 0 & 0 & \upsilon_1 & \cdots & \upsilon_{r-2} & \upsilon_{r-1} & \upsilon_r & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \upsilon_1 & \upsilon_2 & \upsilon_3 & \cdots & \upsilon_r \end{pmatrix}.$$

Suppose now $\upsilon_1 \neq 0$, otherwise, take υ_2 and so-on and so-forth. We can always find a non-zero entry of υ since it is non-zero. The $(r+1)\times(r+1)$ sub-matrix of ∇g given by

$$\begin{pmatrix} \upsilon_1 & \upsilon_2 & \upsilon_3 & \cdots & \upsilon_r & * \\ 0 & \upsilon_1 & \upsilon_2 & \cdots & \upsilon_{r-1} & * \\ 0 & 0 & \upsilon_1 & \cdots & \upsilon_{r-2} & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \upsilon_1 & * \\ 0 & 0 & 0 & \cdots & 0 & 2\upsilon_1 \end{pmatrix} \in \mathbb{R}^{(r+1)\times(r+1)}$$

of $\nabla g(\eta, \upsilon)$ is upper-diagonal and thus has non-zero determinant $2(\upsilon_1)^{r+1}$. Thus, we have found a non-zero minor of $\nabla g(\eta, \upsilon)$ which is of order $r+1$. The matrix $\nabla g(\eta, \upsilon)$ has $r+1$ rows so there cannot exist any minors of higher order. By Proposition A.1.12 it then holds that $\operatorname{rank} \nabla g(\eta, \upsilon) = r+1$. □

4.2.3 The Riemannian SVD

We fix now an affine matrix-function $\mathcal{B} : \mathbb{R}^n \to \mathbb{R}^{p\times m}$ and a vector $a \in \mathbb{R}^n$.

In [Moo94], the author introduces the following set of equations called the **Riemannian SVD**:

$$\begin{cases} \mathcal{B}(a)v = \tau D_v u, & u^\top D_v u = 1 \\ \mathcal{B}(a)^\top u = \tau D_u v, & v^\top D_u v = 1, \end{cases} \tag{RSVD}$$

where

$$D_u := \sum_{i=1}^{n} B_i^\top (u^\top B_i(\cdot))u,$$
$$D_v := \sum_{i=1}^{n} B_i((\cdot)^\top B_i v)v,$$

and shows that they play an important role in describing the solutions to STLS problems in general and HTLS in particular.

Before presenting the main result, we note that D_u is quadratic in the components of u. For $x = \sigma^{-1}u$, where $\sigma = \|u\|_2$, we have that

$$D_u = \sum_{i=1}^{n} B_i^\top (u^\top B_i(\cdot))u = \sum_{i=1}^{n} B_i^\top ((\sigma x)^\top B_i(\cdot))\sigma x = \sigma^2 \sum_{i=1}^{n} B_i^\top (u^\top B_i(\cdot))u = \sigma^2 D_x.$$

Similar computation shows that D_v is quadratic in the components of v. Another interesting property of D_u and D_v is that they, for every u and v respectively, are symmetric, positive semi-definite matrices. For example, we have for arbitrary u, v that

$$v^\top D_u v = \sum_{i=1}^{n} v^\top B_i^\top (u^\top B_i v)u = \sum_{i=1}^{n} (u^\top B_i v)^2 \geq 0,$$

since the transpose of a scalar is itself.

With that out of the way, we now formulate the main theorem of [Moo94] in its full generality.

Theorem 4.2.9. *Consider the program to*

$$\text{minimize} \quad \textit{over } b \textit{ and } y \quad \sum_{i=1}^{n} (a_i - b_i)^2 \quad \text{subject to} \quad \begin{cases} b \in \mathbb{R}^n, \\ y \in \mathbb{R}^m, \\ \mathcal{B}(b)y = \vec{0}, \\ y^\top y = 1. \end{cases}$$

A solution (b^*, y^*) to the program is obtained by finding a triple $(u, \tau, v) \in \mathbb{R}^p \times \mathbb{R} \times \mathbb{R}^m$ corresponding to the smallest $|\tau|$ that satisfy the RSVD and setting

$$y^* = \frac{v}{\|v\|_2}, \quad \textit{and} \quad b_i^* = a_i - u^\top B_i v \tau \quad \textit{for } i = 1, 2, \ldots, n.$$

We need the following statement before presenting the proof of the theorem for $\mathcal{S} = \mathcal{H}_n$.

Lemma 4.2.10. If (u_1, τ_1, v_1) solves RSVD, then there are non-zero constants $\alpha_1, \beta_1 \in \mathbb{R}$ such that (u_1', τ_1', v_1'), where

$$u_1' = \alpha_1^{-1} u_1, \quad \tau_1' = \tau_1 \alpha_1 \beta_1, \quad v_1' = \beta_1^{-1} v_1',$$

solves the **Multiplier condition**

$$\begin{cases} \mathcal{B}(a)v = \tau D_v u, & u^\top u = 1 \\ \mathcal{B}(a)^\top u = \tau D_u v, & v^\top v = 1. \end{cases} \tag{MC}$$

Similarly, if (u_2, τ_2, v_2) solves MC, then there are non-zero constants $\alpha_2, \beta_2 \in \mathbb{R}$ such that (u_2', τ_2', v_2'), where

$$u_2' = \alpha_2^{-1} u_2, \quad \tau_2' = \tau_2 \alpha_2 \beta_2, \quad v_2' = \beta_2^{-1} v_2',$$

solves RSVD.

Proof. Suppose the triple (u, τ, v) solves the RSVD. By $u^\top D_v u = v^\top D_u v = 1$, u and v must be non-zero. By setting $u' = \|u\|_2^{-1} u$, $v' = \|v\|_2^{-1} v$ and $\tau' = \tau \|v\|_2 \|u\|_2$, we have that

$$\begin{aligned} \mathcal{B}(a)v' = \mathcal{B}(a)\frac{v}{\|v\|_2} &= (\tau \|v\|_2 \|u\|_2)\frac{1}{\|v\|_2^2} D_v \frac{u}{\|u\|_2} = \tau' D_{v'} u', \\ \mathcal{B}(a)^\top u' = \mathcal{B}(a)\frac{u}{\|u\|_2} &= (\tau \|v\|_2 \|u\|_2)\frac{1}{\|u\|_2^2} D_u \frac{v}{\|v\|_2} = \tau' D_{u'} v'. \end{aligned}$$

Since we normalized, we have that $(u')^\top u' = (v')^\top v = 1$, so (u', τ', v') solves MC.
Suppose (x, σ, y) solves MC. Setting $x' = \alpha^{-1} x$, $y' = \beta^{-1} y$ and $\sigma' = \sigma \alpha \beta$ for some non-zero $\alpha, \beta \in \mathbb{R}$, we have that

$$\begin{aligned} \mathcal{B}(a)(\beta y') = \sigma D_{y'} \beta^2 \alpha x' &\iff \mathcal{B}(a) y' = \sigma' D_{y'} x' \\ \mathcal{B}(a)^\top (\alpha x') = \sigma D_{x'} \alpha^2 \beta y' &\iff \mathcal{B}(a)^\top x' = \sigma' D_{x'} y'. \end{aligned}$$

Note for all (υ, ω) that

$$\upsilon^\top \mathcal{B}(a) \omega = \omega^\top \mathcal{B}(a)^\top \upsilon = \omega^\top D_\upsilon \omega = \upsilon^\top D_\omega \upsilon \geq 0,$$

since D_ω and D_υ are positive semi-definite. We set $\gamma^2 = x^\top D_y x = y^\top D_x y > 0$, from which We get that

$$(x')^\top D_{y'} x' = (y')^\top D_{x'} y' = \gamma^2 (\alpha\beta)^{-2}.$$

Thus, we are done by choosing α, β satisfying $\gamma^2 = \alpha^2 \beta^2$. The reason for γ^2 not being zero is that if it were, then both non-zero vectors x and y would be in the null-space of $\mathcal{B}(a)$, but it is assumed to be of full-rank, so it can't be the case. Thus, we have that (x', σ', y') solves RSVD. ☐

Remark 4.2.11. Thus, we see that the MC and RSVD are in some loose sense equivalent, since each solution for one problem yields a solution for the other.

Proof of Theorem 4.2.9 ($\mathcal{B} = \mathcal{H}_n$). The Lagrangian for HTLS is

$$\mathcal{L}(b, y, \lambda, \mu) = \|a - b\|_2^2 - \lambda^\top \mathcal{H}_n(b) y + \mu(1 - y^\top y).$$

We search for stationary points by setting $\nabla\mathcal{L} = \vec{0}$, which yield the following equations

$$\begin{aligned}
\nabla_b\mathcal{L} &= -2(a-b) - X(\lambda, y) = \vec{0}, \\
\nabla_y\mathcal{L} &= -\mathcal{H}_n(b)^\top\lambda - 2\mu y = \vec{0}, \\
\nabla_\lambda\mathcal{L} &= \mathcal{H}_n(b)y = \vec{0}, \\
\nabla_\mu\mathcal{L} &= y^\top y - 1 = 0,
\end{aligned}$$

where X is defined entry-wise by

$$(X(\lambda, y))_i = \lambda^\top H_i y = \sum_{k+l=i+1} \lambda_k y_l, \quad \text{for } i = 1, 2, \dots, n.$$

Moving terms around gives us the following:

$$\begin{aligned}
2(b-a) &= -X(\lambda, y), \\
\mathcal{H}_n(b)^\top\lambda &= -2\mu y, \\
\mathcal{H}_n(b)y &= \vec{0} \\
y^\top y &= 1.
\end{aligned}$$

Since the LICQ is satisfied throughout the feasible region, we have that the above equations do have at least one solution, since we have found that HTLS has a solution.

From the second equation, we get that

$$0 = \vec{0}^\top\lambda = y^\top\mathcal{H}_n(b)^\top\lambda = -2\mu y^\top y = -2\mu,$$

meaning $\mu = 0$.

We substitute λ with $\gamma = \frac{1}{2}\lambda$ and consider the first equation entry-wise:

$$a_i - b_i = \gamma^\top H_i y, \quad \text{where } i = 1, 2, \dots, n.$$

With this we get that

$$\begin{aligned}
\mathcal{H}_n(a)y &= \left(\sum_{i=1}^n a_i H_i\right) y \\
&= \left(\sum_{i=1}^n (b_i + \gamma^\top H_i y) H_i\right) y \\
&= \left(\sum_{i=1}^n b_i H_i + (\gamma^\top H_i y) H_i\right) y \\
&= \mathcal{H}_n(b)y + \sum_{i=1}^n H_i(\gamma^\top H_i y)y \\
&= D_y\gamma.
\end{aligned}$$

Since we have found that $\mathcal{H}_n(b)^\top\gamma = \vec{0}$, we get similarly that

$$\mathcal{H}_n(a)^\top\gamma = \sum_{i=1}^n H_i^\top(\gamma^\top H_i y)\gamma = D_\gamma y.$$

We now normalize by defining $x = \sigma^{-1}\gamma$ where $\sigma = \|\gamma\|$, which we can do since γ has to be non-zero. If that were not the case, we would have that $(a, *)$ would be a candidate for minimization, but $\mathcal{H}_n(a)$ is assumed to be non-singular, and so $(a, *)$ is not in the feasible set.

We have then that

$$D_x y = \sum_{i=1}^{n} H_i^\top (x^\top H_i y) x = \mathcal{H}_n(a)^\top \gamma \sigma^{-1}$$

and that

$$D_y x = \sum_{i=1}^{n} H_i(((\sigma^{-1}x)^\top H_i y) y = \sigma^{-1}\mathcal{H}_n(a) y,$$

meaning we have the derived the MC for the problem:

$$\begin{aligned} \mathcal{H}_n(a) y &= \sigma D_y x, \quad x^\top x = 1, \\ \mathcal{H}_n(a)^\top x &= \sigma D_x y, \quad y^\top y = 1. \end{aligned}$$

We have already shown that this set of equations, are equivalent to the RSVD in the sense described above. From above, we have that

$$x^\top D_y x \sigma = x^\top \mathcal{H}_n(a) y = y^\top \mathcal{H}_n(a)^\top x = y^\top D_x y \sigma.$$

With this, we can reinterpret the objective function:

$$\begin{aligned} \|a - b\|_2^2 &= \sum_{i=1}^{n} (\gamma^\top H_i y)^2 \\ &= \sum_{i=1}^{n} (x^\top H_i y)^2 \sigma^2 \\ &= x^\top \sum_{i=1}^{n} H_i (x^\top H_i y) y \sigma^2 = x^\top D_y x \sigma^2 = x^\top \mathcal{H}_n(a) y \sigma \end{aligned}$$

If now (u, τ, v) solves the RSVD for HTLS, we have that

$$u^\top \mathcal{H}_n(a) v = \tau (u^\top D_v u) = \tau,$$

and by setting $x = \|u\|_2^{-1} u$, $y = \|v\|_2^{-1} v$ and $\sigma = \tau \|u\|_2 \|v\|_2$ we get a solution for this problem's MC. Using the above objective value formula with (x, σ, y), we get that

$$x^\top \mathcal{H}_n(a) y \sigma = \frac{u^\top}{\|u\|_2} \mathcal{H}_n(a) \frac{v}{\|v\|_2} (\tau \|u\|_2 \|v\|_2) = \frac{u^\top D_v u}{\|u\|_2 \|v\|_2^2 \|u\|_2} (\tau \|u\|_2 \|v\|_2)^2 = \tau^2.$$

Hence, we see that the key is to find a triple (u, τ, v) corresponding to the smallest $|\tau|$. We see here also that $y = \|v\|_2^{-1} v$, and we get for $k = 1, 2, \ldots, n$ that

$$b_k = a_k - x^\top H_i y \sigma = a_k - u^\top H_i v \tau,$$

since $\gamma = \sigma x$. □

Remark 4.2.12. The solution to the HTLS is not unique since the vectors u and v are not unique to every τ that together solve the RSVD. This can be seen in the following way. Let (u, τ, v) solve RSVD and let $\alpha, \beta \in \mathbb{R}$ such that $\alpha\beta = 1$. By setting $u' = \alpha^{-1}u$ and $v' = \beta^{-1}v$, we have that

$$\mathcal{B}(a)v' = \mathcal{B}(a)\frac{v}{\beta} = (\tau\alpha\beta)\frac{1}{\beta^2}D_v\frac{u}{\alpha} = \tau D_{v'}u',$$
$$\mathcal{B}(a)^\top u' = \mathcal{B}(a)\frac{u}{\alpha} = (\tau\alpha\beta)\frac{1}{\alpha^2}D_u\frac{v}{\beta} = \tau D_{u'}v',$$

and that

$$(u')^\top D_{v'}u' = u^\top D_v u = 1,$$
$$(v')^\top D_{v'}u' = v^\top D_u v = 1,$$

implying that (u', τ, v') also solves the RSVD.

The problem of actually solving RSVD is beyond the scope of this thesis, but the interested reader can find an algorithm on how to numerically solve them in [Moo94].

We end this chapter with a necessary condition which might be helpful in finding candidates for minimization of the HTLS.

Corollary 9. If (b, y) solves the HTLS, then

$$(a - b)^\top b = 0,$$

i.e., the vectors b and $(a - b)$ are orthogonal.

Proof. If (b, y) solves HTLS, we have from the proof of the previous theorem that there is a non-zero vector γ such that

$$\gamma^\top \mathcal{H}_n(b)y = 0,$$

and that the components of $a - b$ are given by

$$\gamma^\top H_i y, \quad \text{for } i = 1, 2, \ldots, n.$$

Thus, we have that

$$0 = \gamma^\top \mathcal{H}_n(b)y = \sum_{i=1}^{2n-1} b_i \gamma^\top H_i y = \sum_{i=1}^{2n-1} b_i(a_i - b_i) = b^\top(a - b).$$

□

Chapter 5

Conclusion and Further Reading

The aim of this thesis was to consider the problems of

1. realization,
2. partial realization,
3. approximate realization,

and provide solution methods for each of them, as well as think about how large/ small the resulting models were. In this section, we recap some of the most important moments, comment on some questions that were not treated in this thesis, as well as guide the reader to alternative treatments of the problems that were not mentioned.

5.1 The Partial Realization Problem

The main properties that differentiate the problem of partial realization from the problem of realization is the fact that solutions always exist, and that minimality behaves differently.

In Chapter 3, we investigated under which conditions our partial realizations are in fact minimal, and presented an algorithm that turns any partial realization into a minimal one. The question of minimality was only partly addressed towards the end of the chapter, where we proved minimality under the very restrictive condition $\alpha + \beta \leq r$. In general, minimal partial realization are not unique, so it is of interest to study in which ways we can represent all minimal realizations with a single set of formulas. The parameterization of all minimal realizations for a truncated Markov sequence is as of 1990 still not solved according to [BGR90]. Thus, it could be of interest to see if a parameterization has been discovered since then.

We presented a few ways to construct partial realizations, one of which is via re-interpretation of the problem to one of rational interpolation. This was only done for the SISO case where we have an odd number of Markov parameters. In [AA90] and in [AMY91], this restriction is nowhere to be seen, meaning that one should be able to re-formulate and solve the partial realization problem for the general $p \times m-$case. What is not mentioned is that, for the re-interpretation and subsequent solving of the partial realization problem via rational interpolation to make sense, we needs to identify the conditions for rational matrix function $W(1/s)$ to be proper provided we have a rational matrix-valued function W. Furthermore, authors of [AMY91] state that the

McMillan degree of $W(1/s)$ and $W(s)$ are the same, but leaves no references to any proof of the statement. Thus, an investigation into when the McMillan-degree of W remains the same after pre-composition by $s \mapsto s^{-1}$ is warranted if one would like to extend/ apply the methods of this thesis to the matrix-case.

A more recent interpretation of the partial realization problem in the SISO-and-real case is through the rank-minimization of the Hankel matrix. In [FHB04], the problem of minimal partial realization for a truncated real Markov sequence $h = (h_1, h_2, \dots, h_n)^\top$ is formulated as the following problem to

$$\text{minimize over } \hat{h} \quad \operatorname{rank} \mathcal{H}_{n,n}(\hat{h}) \text{ subject to } \hat{h}_i = h_i \text{ for all } i = 1, 2, 3, \dots, n.$$

The motivation behind this reformulation is the fact that the degree of h is can calculated by

$$\min\{\operatorname{rank} \mathcal{H}_{n,n}(\hat{h}) : \hat{h}_i = h_i \text{ for all } i = 1, 2, 3, \dots, n\},$$

which was also shown in the article. Since rank optimization problems are in general difficult, the authors of the article opted to solve a convex surrogate problem with the *Nuclear Norm*. An interesting thing to look at would be when a solution to the surrogate problem is a solution to the rank-optimization problem.

5.2 The Approximate Realization Problem

The approximate realization problem can be viewed as a relaxation of the partial realization problem since we now search for systems which only approximate our input-output data. In Chapter 4, we presented Kung's algorithm, and mentioned some of its features, for example the loss of Hankel structure in using an unstructured low-rank approximation. We then considered ways to construct approximate realizations provided we had access to a low-rank approximation with Hankel structure. Lastly, we showed how to find a low-rank approximation that has Hankel structure. One thing to note is that our approach is in SISO. This is partly due to the fact that in [Moo94], they only dealing with Hankel structure and presenting the problem as minimization of the $2-$norm of the difference of vectors in some $\mathbb{R}^n$. In the general case, we would have needed to deal with block-Hankel structure and minimization of some other type of objective function. Furthermore, the assumption in the HTLS problem is that the Hankel-matrix is full-rank, which in the general matrix case would be hard to work with. Another reason for only considering the SISO-case is that we have found a class of truncated Markov sequences we can use it on.

An alternative way one might be able to go about constructing an approximate realization is by using the characterization and methods in [Ast07] and in [Pad23]. In [Ast07], the author introduced the notion of moments of a transfer function in a point on the plane, which are essentially interpolation conditions. The author then showed that the moments of a transfer function has a one-to-one correspondence to the solution of a certain Sylvester equation. With this characterization, he then went on to characterize what is essentially a class of systems which has prescribed moments, i.e., they satisfy interpolation conditions, and are of low dimension than a system that is known the prescribed moments. In [Pad23], this moment-matching condition is relaxed to be a least-squares moment-matching condition, which gives a way to find systems whose moments are approximately equal to those we want. Both of these article have defined moments to be at points which are not infinity. It could thus be of interest to see if we get approximate realizations of truncated Markov sequences by first considering a realization, whose transfer function W has the truncated Markov sequences as its moments at the origin, and then finding via the model reduction scheme in [Pad23], a system which approximate those moments and then pre-composing by $s \mapsto s^{-1}$ to get $W \circ ()^{-1}$.

Appendix A

Some Useful Results

A.1 Matrix Algebra

Proposition A.1.1. Let $A \in \mathbb{K}^{n\times n}$ be matrix such that

$$\|A\|_F := \sqrt{\sum_{i,j} |A_{i,j}|^2} < 1.$$

Then $I - A$ is non-singular and

$$(I - A)^{-1} = \left(\sum_{i=0}^{\infty} A^k\right) \in \mathbb{K}^{n\times n}.$$

Proof. Apply Theorem 3.7 in [DP00, Chapter 3.2] to the Banach algebra $\mathcal{B} = (\mathbb{K}^{n\times n}, \|\cdot\|_F)$. The fact that this is a Banach algebra be seen by $(\mathbb{K}^{n\times n}, \|\cdot\|_F)$ being a Banach space (since it is finite-dimensional) and $\|\cdot\|_F$ being sub-multiplicative. For a proof of the latter, see [htt]. $\square$

Definition A.1.2. A matrix $A \in \mathbb{K}^{k\times l}$ has

- full column-rank if $\operatorname{rank} A = l$,
- full row-rank if $\operatorname{rank} A = k$.

Proposition A.1.3. Let A be a $k \times l$ matrix.

- If A has full column-rank, then A^*A is non-singular and $A^{-L} := (A^*A)^{-1}A^*$ is the *left pseudo-inverse* of A, meaning $A^{-L}A = I \in \mathbb{K}^{l\times l}$.
- If A has full row-rank, then AA^* is non-singular and $A^{-R} := A^*(AA^*)^{-1}$ is a *right pseudo-inverse* of A, meaning $AA^{-R} = I \in \mathbb{K}^{k\times k}$.

Proof. Suppose A has full column-rank. Since A can be viewed as a linear operator $\mathrm{L}_A : \mathbb{K}^l \to \mathbb{K}^k$ we have by rank-nullity that A is injective, due to $\ker A$ having dimension 0. Take now $x \in \mathbb{K}^l$ such that $A^*Ax = \vec{0}$. We have then that $0 = \langle A^*Ax, x\rangle = \langle Ax, Ax\rangle = \|Ax\|_2^2$, meaning $x = \vec{0}$ since A is injective. Thus, A^*A is injective as a linear operator, meaning it is non-singular. Hence, A^{-L} is well-defined and $A^{-L}A = (A^*A)^{-1}A^*A = I$. A similar argument works for showing that AA^* is non-singular when A has full row-rank. $\square$

Proposition A.1.4. Let $H \in \mathbb{K}^{k\times l}$ be a matrix such that $\operatorname{rank} H = r > 0$. There exists infinitely many matrix-pairs $X \in \mathbb{K}^{k\times r}, Y \in \mathbb{K}^{r\times l}$ where X has full column-rank and Y has full row-rank, such that

$$H = XY.$$

Furthermore, if $H = XY$ with $\operatorname{rank} X = \operatorname{rank} Y = \operatorname{rank} H$, a so-called *full-rank factorization* of H and UV is also a full-rank factorization of H, then the pairs are similar, i.e. there is a $Y \in \mathrm{GL}_r(\mathbb{K})$ such that $U = OT$ and $V = T^{-1}R$.

Proof. [PO99]. □

Proposition A.1.5. Consider complex numbers $q_1, \dots, q_m$, and the polynomial

$$q(s) = q_0 + q_1 s + q_2 s^2 + \dots + q_m s^m.$$

The matrix A, given by

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 1 \\ -q_{m-1} & -q_{m-2} & -q_{m-3} & \cdots & -q_0 \end{pmatrix} \in \mathbb{C}^{m\times m},$$

satisfies

$$\chi_A = \det(sI - A) = q.$$

Proof. [Son98, Chapter 5.1]. □

Proposition A.1.6. Let

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

be a block-matrix with A being $m \times m$ and non-singular and D being $n \times n$ and B and C having proper dimensions for M to be square. Then

$$M^{-1} = \begin{pmatrix} X & Y \\ U & V \end{pmatrix}$$

with

$$\begin{aligned} X &= A^{-1} - A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1}, \\ Y &= -A^{-1}B(D - CA^{-1}B)^{-1}, \\ U &= -(D - CA^{-1}B)^{-1}CA^{-1}, \\ V &= (D - CA^{-1}B)^{-1}. \end{aligned}$$

Proof. [Zha11, Chapter 2.2]. □

Proposition A.1.7. Let

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

be a block-matrix with A being $m \times m$ and non-singular and D being $n \times n$ and B and C having proper dimensions for M to be square. Then

$$\det M = \det A \det(D - CA^{-1}B).$$

Proof. [Zha11, Chapter 2.2]. $\square$

Proposition A.1.8. The matrix equation

$$AX + XB + C = \mathbf{0}$$

admits a unique solution X if and only if the spectra of A and $-B$ are disjoint.

Proof. [Son98, Appendix A.4]. $\square$

Proposition A.1.9. Suppose $A \in \mathbb{K}^{m\times p}$ and $B \in \mathbb{K}^{p\times n}$. Then

$$\operatorname{rank}(AB) = \operatorname{rank} B - \dim\{\mathrm{Im}B \cap \ker A\}.$$

Proof. [Zha11, Chapter 2.3]. $\square$

Definition A.1.10. Let A be a matrix. A sub-matrix of A is a matrix which is obtained by deleting any number of rows and/or columns from A.

Definition A.1.11. Let A be a matrix. A minor is the determinant of a square sub-matrix of A. The order of a minor is given by the number of rows/columns of the sub-matrix.

Proposition A.1.12. The rank of a matrix is given by the maximal order of a non-zero minor.

Proof. [Rob13, Chapter 1.1]. $\square$

A.2 Equality-constrained Optimization

Definition A.2.1. Let $f : \mathbb{R}^n \to \mathbb{R}$, $g \in \mathbb{R}^n \to \mathbb{R}^m$ be C^1-functions and consider the equality constrained optimization problem

$$\min f(x) \quad \text{s.t} \quad g(x) = \vec{0}.$$

We say that the constraint g_i, for $i = 1, \dots, m$ satisfy the *linear independence constraint qualification* (LICQ) at points x^* if

$$\operatorname{rank} \nabla g(x^*) = m,$$

or equivalently, that $\{\nabla_i g(x^*)\}_{i=1}^m$ constitutes a set of linearly independent vectors.

Theorem A.2.2. *Let $f : \mathbb{R}^n \to \mathbb{R}$, $g \in \mathbb{R}^n \to \mathbb{R}^m$ be C^1-functions and consider the equality constrained optimization problem:*

$$\min f(x) \quad s.t \quad g(x) = 0.$$

If $x^* \in \ker g$ is a local extreme point for which the LICQ holds, then there exists a so-called Lagrange multiplier-vector $\lambda^* \in \mathbb{R}^m$ such that

$$\nabla f(x^*) = (\lambda^*)^\top \nabla g(x^*).$$

Proof. [Rob13, Chapter 3.2] $\square$

Remark A.2.3. Consider the *Lagrangian* of the above problem

$$\mathcal{L}(x, \lambda) = f(x) - \lambda^\top g(x).$$

If x^* is a local extreme point and λ^* is a multiplier associated to that point, we have that

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = \nabla f(x^*) - (\lambda^*)^\top \nabla g(x^*) = \vec{0}.$$

Bibliography

[AA90] B.D.O. Anderson and A.C. Antoulas. “Rational interpolation and state-variable realizations”. In: *Linear Algebra and its Applications* 137-138 (1990), pp. 479–509. ISSN: 0024-3795. DOI: https://doi.org/10.1016/0024-3795(90)90140-8. URL: https://www.sciencedirect.com/science/article/pii/0024379590901408.

[AMY91] A. C. Antoulas, T. Matsūo, and Y. Yamamoto. “Linear Deterministic Realization Theory”. In: *Mathematical System Theory: The Influence of R. E. Kalman*. Springer Berlin Heidelberg, 1991, pp. 191–212. ISBN: 978-3-662-08546-2. DOI: 10.1007/978-3-662-08546-2_11. URL: https://doi.org/10.1007/978-3-662-08546-2_11.

[Ast07] A. Astolfi. “MODEL REDUCTION BY MOMENT MATCHING”. In: *IFAC Proceedings Volumes* 40.12 (2007). 7th IFAC Symposium on Nonlinear Control Systems, pp. 577–584. ISSN: 1474-6670. DOI: https://doi.org/10.3182/20070822-3-ZA-2920.00095. URL: https://www.sciencedirect.com/science/article/pii/S1474667016355884.

[BGR88] J. A. Ball, I. Gohberg, and L. Rodman. “Realization and Interpolation of Rational Matrix Functions”. In: *Topics in Interpolation Theory of Rational Matrix-valued Functions*. Birkhäuser Basel, 1988, pp. 1–72. ISBN: 978-3-0348-5469-6. DOI: 10.1007/978-3-0348-5469-6_1. URL: https://doi.org/10.1007/978-3-0348-5469-6_1.

[BGR90] Joseph A. Ball, Israel Gohberg, and Leiba Rodman. “Partial Realization as an Interpolation Problem”. In: *Interpolation of Rational Matrix Functions*. Basel: Birkhäuser Basel, 1990, pp. 371–378. ISBN: 978-3-0348-7709-1. DOI: 10.1007/978-3-0348-7709-1_18. URL: https://doi.org/10.1007/978-3-0348-7709-1_18.

[Bos83] O.H Bosgra. “On parametrizations for the minimal partial realization problem”. In: *Systems & Control Letters* 3.4 (1983), pp. 181–187. DOI: https://doi.org/10.1016/0167-6911(83)90011-7. URL: https://www.sciencedirect.com/science/article/pii/0167691183900117.

[DP00] G. E. Dullerud and F. Paganini. *A Course in Robust Control Theory: A Convex Approach*. Springer New York, NY, 2000. ISBN: 978-0-387-98945-7.

[FHB04] M. Fazel, H. Hindi, and S. Boyd. “Rank minimization and applications in system theory”. In: *Proceedings of the 2004 American Control Conference*. Vol. 4. 2004, 3273–3278 vol.4. DOI: 10.23919/ACC.2004.1384521. URL: https://ieeexplore.ieee.org/document/1384521.

[Gil63] E. G. Gilbert. “Controllability and observability in multivariable control systems”. In: *SIAM Journal on Control* 1.2 (1963), pp. 128–151. DOI: 10.1137/0301009. URL: https://epubs.siam.org/doi/10.1137/0301009.

[GKL87] I. Gohberg, M.A. Kaashoek, and L. Lerer. “On minimality in the partial realization problem”. In: *Systems & Control Letters* 9.2 (1987), pp. 97–104. DOI: 10.1016/0167-6911(87)90015-6. URL: https://www.sciencedirect.com/science/article/pii/0167691187900156.

[GM10] A. Gombani and G. Michaletzky. “On the partial realization problem”. In: *Proceedings of the 19th International Symposium on Mathematical Theory of Networks and Systems*. 2010. URL: https://api.semanticscholar.org/CorpusID:18537878.

[Gra17] M. Grasmair. *Minimizers of optimization problems*. 2017. URL: https://wiki.math.ntnu.no/_media/tma4180/2015v/existence.pdf.

[HK66] B. L. Ho and R. E. Kalman. “Die Konstruktion von linearen Modeilen in der Darstellung durch Zustandsvariable aus den Beziehungen für Ein- und Ausgangsgrößen”. In: *at-Automatisierungstechnik* 14.1-12 (1966), pp. 545–548. DOI: doi:10.1524/auto.1966.14.112.545.

[htt] B E I R U T (https://math.stackexchange.com/users/859737/b-e-i-r-u-t). *Showing that Frobenius Norm is Submultiplicative*. Mathematics Stack Exchange. URL:https://math.stackexchange.com/q/3966454 (version: 2020-12-30). eprint: https://math.stackexchange.com/q/3966454. URL: https://math.stackexchange.com/q/3966454.

[JP85] Jer-Nan Juang and Richard S. Pappa. “An eigensystem realization algorithm for modal parameter identification and model reduction”. In: *Journal of Guidance, Control, and Dynamics* 8.5 (1985), pp. 620–627. DOI: 10.2514/3.20031. eprint: https://doi.org/10.2514/3.20031. URL: https://doi.org/10.2514/3.20031.

[Kal63] R. E. Kalman. “Mathematical description of linear dynamical systems”. In: *SIAM Journal on Control* 1.2 (1963), pp. 152–192. DOI: 10.1137/0301010. URL: https://doi.org/10.1137/0301010.

[Kun78] S. Kung. “A new identification and model reduction algorithm via singular value decomposition”. In: *Proc. 12th asilomar conf. on circuits, systems and computer*. 1978, pp. 705–714.

[Mar11] Ivan Markovsky. *Low Rank Approximation: Algorithms, Implementation, Applications*. Springer London, 2011. ISBN: 978-1-4471-2226-5.

[Moo81] B. Moore. “Principal component analysis in linear systems: Controllability, observability, and model reduction”. In: *IEEE Transactions on Automatic Control* 26.1 (1981), pp. 17–32. DOI: 10.1109/TAC.1981.1102568.

[Moo94] B. De Moor. “Total least squares for affinely structured matrices and the noisy realization problem”. In: *IEEE Transactions on Signal Processing* 42.11 (1994), pp. 3104–3113. DOI: 10.1109/78.330370.

[nd] *Systemteori - Övning 4: Realization Theory*. n.d. URL: https://www.math.kth.se/optsyst/grundutbildning/kurser/SF2832/Notes/ovning4.pdf.

[Pad23] A. Padoan. “Model Reduction by Least Squares Moment Matching for Linear and Nonlinear Systems”. In: *IEEE Transactions on Automatic Control* 68.12 (2023), pp. 8267–8274. DOI: 10.1109/TAC.2023.3294869.

[PCM86] G. Pitstick, J. Cruz, and R. Mulholland. “Approximate realization algorithms for truncated impulse response data”. In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 34.6 (1986), pp. 1583–1588. DOI: 10.1109/TASSP.1986.1164997.

[PO99] R. Piziak and P. L. Odell. "Full Rank Factorization of Matrices". In: *Mathematics Magazine* 72.3 (1999), pp. 193–201. DOI: 10.1080/0025570X.1999.11996730. URL: https://www.jstor.org/stable/2690882.

[Ris71] J. J. Rissanen. "Recursive Identification of Linear Systems". In: *SIAM Journal on Control* 9.3 (1971), pp. 420–430. DOI: 10.1137/0309031. URL: https://doi.org/10.1137/0309031.

[Rob13] C. Robinson. *Introduction to mathematical optimization*. Department of Mathematics, Northwestern University, Illinois US, 2013.

[Sch00] B. De Schutter. "Minimal state-space realization in linear system theory: an overview". In: *Journal of Computational and Applied Mathematics* 121.1 (2000), pp. 331–354. DOI: https://doi.org/10.1016/S0377-0427(00)00341-1. URL: https://www.sciencedirect.com/science/article/pii/S0377042700003411.

[Son98] E. D. Sontag. *Mathematical Control Theory: Deterministic Finite Dimensional Systems*. Springer New York, NY, 1998. ISBN: 978-0-387-98489-6.

[Zha11] F. Zhang. *Matrix Theory: Basic Results and Techniques*. Springer New York, NY, 2011. ISBN: 978-1-4614-1098-0.

Errata for *State-Space Realization and Model-Order Reduction: Modeling under Complexity Constraints*
Shamiur Rahman Ramim
Spring 2026

(1) p.13: "...for all admissable t."

The number t is here assumed to be real.

(2) p.15: "..., is it really the case that

$$W = W_\Sigma$$

everywhere on the complex plane?"

The unspecified system Σ is given by $\Sigma = (A, B, C, W(\infty))$.

(3) p.47: "...

$$= \sum_{k=0}^{r-1} TF^k G u_k$$
$$= \sum_{k=0}^{r-1} T \begin{pmatrix} \mathbf{0} \\ F_{00}^k G_0 \end{pmatrix} u_j = T \begin{pmatrix} \vec{0} \\ \sum_{k=0}^{r-1} F_{00}^k G_0 u_k \end{pmatrix}"$$

It should be u_k and not u_j.

(4) p.49: "...

$$= \sum_{k=0}^{r-j-1} TF^k G u_k$$
$$= \sum_{k=0}^{r-j-1} T \begin{pmatrix} \mathbf{0} \\ F_{00}^k G_0 \end{pmatrix} u_j = T \begin{pmatrix} \vec{0} \\ \sum_{k=0}^{r-j-1} F_{00}^k G_0 u_k \end{pmatrix}"$$

It should be u_k and not u_j.

(5) p.59: In Lemma 4.1.2, D is supposed to be a square $p \times p$ real matrix and m a positive integer no greater than p.

In Example 14, we have assumed that the singular values $\sigma_1, \sigma_2, \ldots, \sigma_m$ are all non-zero, for $m < p$.

In Example 14, it should be $O = U_1 \Sigma_1^{1/2}$ and $R = \Sigma_1^{1/2} V_1^\top$ instead of $O = U_1 \Sigma^{1/2}$ and $R = \Sigma^{1/2} V_1^\top$.

In Example 14, it is stated that O_f, which is the matrix made up of the first rows of O is full column rank since O is. This is incorrect, meaning we cannot in general get the matrix A from the equation $O_f A = O_l$ by multiplying with $O_f^\top$ to then multiplying by $(O_f^\top O_f)^{-1}$ get $A = (O_f^\top O_f)^{-1} O_f^\top O_l$. In cases where O_f is rank deficient, we may instead want to construct A by picking one that minimizes $\|O_f A - O_l\|_F$.

In Example 14, it is stated that O_f is made up of the "first $n-1$ vectors of O", but the correct statement is that O_f is made up of the "first $r-1$ rows of O".

(6) p.63: "By linearity of $\mathcal{H}$, ..."

It is supposed to be $\mathcal{H}_r$, which is the (r, r)th Hankel affine matrix-function.

"...

$$\mathcal{F} = \{(b, y) \in \mathbb{R} \times \mathbb{R} : \mathcal{H}_1(b) = by = 0,\ y^2 = 1\} = \{(0, 1), (0, -1)\}."$$

It is supposed to be "$H_1(b)y = by$".

(7) p.66: "We fix now an affine matrix-function $\mathcal{B} : \mathbb{R}^n \to \mathbb{R}^{p\times m}$ and a vector $a \in \mathbb{R}^n$."

The vector a is chosen among the vectors that make $\mathcal{B}(a)$ have full column-rank.

"We need the following statement before presenting the proof of the theorem for $\mathcal{S} = \mathcal{H}_n$."

It is supposed to be "$\mathcal{B} = \mathcal{H}_n$".

(8) p.67: "Note for all (υ, ω) that

$$\upsilon^\top \mathcal{B}(a)\omega = \omega^\top \mathcal{B}(a)^\top \upsilon = \omega^\top D_\upsilon \omega = \upsilon^\top D_\omega \upsilon \geq 0,$$

..."

This is incorrect. What it is meant to be is that every triple $(\upsilon, \sigma, \omega)$ that solves the MC (or the RSVD) satisfies

$$\upsilon^\top \mathcal{B}(a)\omega = \omega^\top \mathcal{B}(a)^\top \upsilon = \sigma\omega^\top D_\upsilon \omega = \sigma\upsilon^\top D_\omega \upsilon,$$

and that $\omega^\top D_\upsilon \omega = \upsilon^\top D_\omega \upsilon$ are is non-negative since D_υ and D_ω are positive semi-definite.

"..., then both non-zero vectors x and y would be in the null-space of $\mathcal{B}(a)$, ..."

It is correct that y would lie in the null-space of $\mathcal{B}(a)$, but x would instead lie in the null-space of $\mathcal{B}(a)^\top$.

(9) p.71: "The question of minimality was only partly addressed towards the end of the chapter, where we proved minimality under the very restrictive condition $\alpha + \beta \leq r$."

It should be "uniqueness of minimal realizations" and not just "minimality".