ADDIS ABABA UNIVERSITY
COLLEGE OF NATURAL SCIENCES
DEPARTMENT OF MATHEMATICS

OBSERVABILITY OF LINEAR TIME INVARIANT DYNAMICAL SYSTEM

A Project Submitted to the Department of Mathematics of Addis Ababa University in Partial Fulfillment of the Requirements of the Master of Science Degree in Mathematics

Abdurahman Jundi Sheko
Advisor: Berhanu B. (Ph.D)

June, 2017

Addis Ababa University, Department of Mathematics

*We,the undersigned, hereby certify that we have read and examined this project paper, a project on **Observability of linear time invariant dynamical systems**, which is done by Abdurahman Jundi Sheko in partial fulfillment of the requirements for the degree of master of science and recommend to the school of graduate studies for acceptance of the project.*

Advisor: Dr. Berhanu Bekele

Signature:______________________________

Date______________________________

Examiner

1. Dr. ______________________________

Signature:______________________________

Date______________________________

2. Dr. ______________________________

Signature:______________________________

Date______________________________

Contents

Acknowledgements

God gave you a gift of 86,400
seconds to day. have you used
one to say 'thank you'?
-*William Arthur Ward*
(1921 – 1994)

I would like to express my special appreciation and thanks to my advisor Dr. Berhanu Bekele. Dr- Berhanu you have been a tremendous mentor for me! You have shown the attitude and the substance of a genius: you continually and persuasively conveyed a sprit of adventure in regard to the project.you were constantly encouraging my work and also allowed me to find my own track as mathematician. you taught me how to deliver the best talk, you have spent so much time listening to my practice on seminar you prepared for us suggesting improvements and offering career advise. All you did for me has been priceless. With you everything, indeed, is fun and a joy and hence my work has become almost a leisure activity! I would also like to thank all Mathematics department teachers for their helpfulness and willingness to chat about my work. Their advise was very helpful in my formation as Mathematician.
The following people deserve special thanks. Their continuous chat on my work initiated and helped to grow my desire as mathematician. I am sure that without them this work would have been impossible. Your support and constant interest in my progress and achievement during my presentation on seminar were very encouraging!

1. *Yibeltal Y. (PHD)*
2. *Tilahun A. (PHD)*
3. *Zelalem T. (PHD)*

I am filled with gratitude for many interesting discussion and chats with PHD students of the year and MSc student Meseret Abiy an Wosenu Abate. I, specially, feel indebted to my colleague Asmamaw Msgan for his unbounded encouragement and support in making my work successful.
Furthermore, I would also like to thank my parents for their endless love and support. Finally, and most importantly to me, I would like to express appreciation to my beloved wife Asiya Mohammed who spent many sleepless night, especially during my stay in library, and was always my support in the moment when their was no one to answer my queries. My love, I am sure, I would not have made it thus far! Your help, Support, and your faith in me was the fuel which makes me go. I will be grateful forever for your love. Thank you!

Abstract

The universe is governed by
science. But science tells us that
we can't solve the equations
directly in the abstract.
-*Stephen Hawking*(1942)

This project paper is aimed to explain observability of linear time invariant dynamical system. We develop a linear systems theory that coincides with the existing theories for continuous and discrete dynamical systems. We explore observability in terms of both Gramian and rank conditions and establish related realizability results.

An observable system is one in which the latent variables can be reconstructed from the manifest variables (in sate space system, the manifest variables are input and output and the latent variable is the state). In order to reconstruct the state at any time from the input and the output, due to the property of state, it suffices to reconstruct state at a specific time t_o, then, we know it every where in the future, i.e, for all $t \geq t_0$. Thus, we only need to reconstruct $x(0)$. We will also state necessary and sufficient conditions for the recostructiblity of the state $x(0)$ or observability of the system, namely, ***Kalman*** observability test, ***Hautus*** observability test and observability test using the ***Gramian*** matrix of the system.

In addition, if the system is not observable, i.e, if the state $x(0)$ is not reconstructible, using ***Kalman*** observability decomposition, we will identify which components of $x(0)$ are reconstructible and which are not. Finally we will give a test for observability of a behaviour. Some examples are included to show the utility of these results.

The first chapter of the paper mainly discusses basic preliminaries for the discussion of the main topic "Observability of linear time invariant dynamical system". In here we will define several terminologies both verbally and mathematically. we will also study and proof some basic theorems. The later chapter discusses observability for linear time invariant dynamical system. Several system properties will be developed and used in checking observability of a given system and proofing system related theorems.

Chapter 1

INTRODUCTION

Everything comes if
you will only wait.
Tancred.

1.1 What is a system?

The word "system" comes from the Greek word $\sigma\upsilon\sigma\tau\eta\mu\alpha$, which originally meant something like "to stand/put/place together."
The following are some definitions i got from some books:

- Systems, in one sense, are devices that take input and produce an output. A system can be thought to operate on the input to produce the output. The output is related to the input by a certain relationship known as the system response. The system response usually can be modeled with a mathematical relationship between the system input and the system output.(*Wikibooks.org*)
- System is a thing that has components which may act independently, but are connected somehow. It is a collection of parts that interact with each other to function as a whole.(*Eva Zerz: Introduction to systems and control theory)*

Physical systems can be divided up into a number of different categories, depending on particular properties that the system exhibits. Some of these system classifications are very easy to work with and have a large theory base for analysis. Some system classifications are very complex and have still not been investigated with any degree of success. By properly identifying the properties of a system, certain analysis and design tools can be selected for use with the system.

Linear time-invariant (LTI) systems are the easiest class of system to work with, and have a number of properties that make them ideal to study. In this chapter we will discuss some properties of this system.

Based on memory we can classify systems as **Dynamic** and **static** systems. A system is said to have memory if the output from the system is dependent on past inputs (or future inputs!) to the system. A system is called memoryless if the output is only dependent on the current input. Memoryless systems are easier to work with, but systems with memory are more common in digital signal processing applications. Systems that have memory are called **dynamic systems**, and systems that do not have memory are **static systems**.

Static system is a system which is fixed with time. It is a system that remains unchanged when a particular transformation is applied to it.

Dynamical system is a system that evolve in time. Such systems interact with their environment through time dependent functions called signals or trajectories.
In here we are interested only on dynamical system. This system is a system in which components/parts/elements evolve in time. According to J.C.Willems' definition of dynamical system (1986/87), a dynamical system Σ involves three ingredients.

- A set $\mathbf{T}$: It is our mathematical model of time. This time set can be either discrete or continuous time.
- Signal value set $\mathbf{W}$: It is a set in which the signals take their values. Thus, a signal is a function

$$w : \mathbf{T} \to \mathbf{W}, t \mapsto w(t).$$

- Behavior $\mathcal{B} \subseteq \mathbf{W}^T$, where $\mathbf{W}^T$ is the set of all functions defined on $\mathbf{T}$ and take their value in $\mathbf{W}$. It contains signals that are compatible with the law governing the system(usually this laws are of the form difference or differential equation). Such signals that obey the system law under consideration are also called admissible signals.

$$\mathcal{B} = \{w \in \mathbf{W}^{\mathbf{T}} | w \text{ satisfies the system law}\}$$

Thus, Willems' famous definition of a dynamical system Σ is a triplet:

$$\Sigma = (\mathbf{T}, \mathbf{W}, \mathcal{B})$$

In most cases of relevance,we consider $\mathbf{T}$ as subset of $\mathbb{R}$ and $\mathbf{W}$ as a finite dimensional vector space, say $\mathbf{W} = \mathbb{K}^q$, where $\mathbb{K}$ is afield under consideration(often $\mathbb{R}$ or $\mathbb{C}$).
Obrest,slightly modified the above definition of willems' but still with similar meaning in 1990 as follow. Let

- $\mathcal{A}$ denotes the set of scalar valued signals

- q denotes the number of signals occurring in a system
- $\mathcal{B} \subseteq \mathcal{A}^q$ denotes the set of q-tuples of such signals that obey the system law

Then, a dynamical system Σ is determined by these three details.

$$\Sigma = (\mathcal{A}, q, \mathcal{B})$$

This Obrest definition is modified version of Willems' definition and more power full as it encompasses distributions.

Remark 1.1. *Prototypes of signal sets: If our signals set is*

- *the set of all functions from $\mathbb{N}$ to $\mathbb{R}$, then such functions are usually called sequences and in this case we take*

$$\mathcal{A} = \mathbb{R}^{\mathbb{N}}$$

- *the set of k times continuously differentiable functions from $\mathbb{R}$ to $\mathbb{R}$ (where $0 \leq k \leq \infty$), we take*

$$\mathcal{A} = \mathbf{C}^k(\mathbb{R})$$

- *The set of generalized functions or distributions, we take $\mathcal{A} = \mathfrak{D}'(\mathbb{R})$*

Example 1.1.1. *The motion of two planets around the sun is governed by Kepler's laws. The time set is certainly continuous. Whether you choose $\mathbb{R}$ or $\mathbb{R}_+$ as your time model, depends on your religious and/or scientific beliefs. Let's not touch these delicate issues and choose $\mathbf{T} = \mathbb{R}$ for simplicity. The position of a planet in space is determined by its three real coordinates, therefore $K = \mathbb{R}$ and $q = 6$, or $\mathbf{W} = \mathbb{R}^6$. If we let $\mathcal{A}$ be the set of all functions from $\mathbb{R}$ to $\mathbb{R}$, then $\mathcal{A}^q = \mathbf{W}^{\mathbf{T}}$. Being inhabitants of the Earth, we may suspect that $\mathcal{A} = \mathbf{C}^k(\mathbb{R})$ for some k which is large enough for comfort. Anyhow, we put*

$$\mathcal{B} = \{w \in \mathcal{A}^6 | w \textit{ satisfies Kepler's laws}\}.$$

Example 1.1.2. ***Genetics/Gender-linked genes**: An allele (a certain form of a gene) is located on the **X**-chromosome. Females have two **X**-chromosomes, males have one **X**- and one **Y**-chromosome. Let $p^f(i)$ be the frequency of the allele in the female gene pool of the i-th generation and let $p^m(i)$ be the same for the male gene pool. Since a son inherits his **X**-chromosome from the mother,*

$$p^m(i+1) = p^f(i)$$

*and since a daughter receives one**X**-chromosome from the father and one from the mother,*

$$p^f(i+1) = \tfrac{1}{2}(p^f(i) + p^m(i)).$$

The time set is discrete. Both $\mathbf{T} = \mathbb{Z}$ and $\mathbf{T} = \mathbb{N}$ are suitable choices. We have two signals p^m and p^f , taking their values in $\mathbb{R}$, hence $\mathbf{W} = \mathbb{R}^2$ and

$$\mathcal{B} = \{ \begin{bmatrix} p^m \\ p^f \end{bmatrix} \in \mathbf{W}^T | \begin{bmatrix} p^m(i+1) \\ p^m(i+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} p^m(i) \\ p^f(i) \end{bmatrix}, \forall i \in \mathbf{T}\}$$

1.2 System properties :LTID systems

For the definition of linearity, we need to fix an underlying number field $\mathbf{K}$. We will focus on $\mathbf{K} = \mathbb{R}$, and therefore we give the definition for that case only. A system is considered linear if it satisfies the conditions of Additivity and Homogeneity.In short, a system is linear if the following is true:

Take two arbitrary inputs, and produce two arbitrary outputs:

$$y_1 = f(x_1)$$
$$y_2 = f(x_2)$$

Now, a linear combination of the inputs should produce a linear combination of the outputs: $f(\lambda_1 x_1 + \lambda_2 x_2) = f(\lambda_1 x_1) + f(\lambda_2 x_2) = \lambda_1 f(x_1) + \lambda_2 f(x_2)$. This condition of additivity and homogeneity is called superposition. A system is linear if it satisfies the condition of superposition.

Definition 1.2.1. *A dynamical system $\Sigma = (\mathcal{A}, q, \mathcal{B})$ is called **linear** if $\mathcal{A}$ is a real vector space, and $\mathcal{B}$ is a subspace of $\mathcal{A}^q$.*

The first requirement means that linear combinations of signals are again signals,

$$a_1, a_2 \in \mathcal{A}; \lambda_1, \lambda_2 \in \mathbb{R} \Rightarrow \lambda_1 a_1 + \lambda_2 a_2 \in \mathcal{A}$$

and the second requirement means that linear combinations of admissible signal vectors are again admissible signal vectors,

$$w_1, w_2 \in \mathcal{B}, \lambda_1, \lambda_2 \in \mathbb{R} \Rightarrow \lambda_1 w_1 + \lambda_2 w_2 \in B.$$

*We call this condition the **superposition principle**.*

Definition 1.2.2 (Shift invariant Dynamical system)**.** *Let $\mathbf{T}$ be such that*

$$t_1, t_2 \in \mathbf{T} \Rightarrow t_1 + t_2 \in \mathbf{T}.$$

For $\tau \in \mathbf{T}$, we define the shift operator σ_τ by

$$\sigma_\tau : \mathbf{W}^{\mathbf{T}} \to \mathbf{W}^{\mathbf{T}}, w \mapsto \sigma_\tau w$$

where

$$(\sigma_\tau w)(t) = w(t + \tau)$$

A dynamical system $\Sigma = (\mathbf{T}, \mathbf{W}, \mathcal{B})$ is called shift-invariant (or: time-invariant) if for all $\tau \in \mathbf{T}$

$$w \in \mathcal{B} \Rightarrow \sigma_\tau w \in \mathcal{B}.$$

Definition 1.2.3. *A dynamical system $\Sigma = (\mathbf{T}; \mathbf{W}; \mathcal{B})$ is called a* ***differential (difference) system*** *if its time set is continuous (discrete) and its system law is given by differential (difference) equations.*

Thus, the class of systems we are going to deal with, in this project, are the class of linear time invariant differential/difference systems, whose system laws are given below:

Differential systems: Systems of linear differential equations with constant coefficients. These can be put in the form

$$(R_d \tfrac{d^d}{dt^d} + ... + R_1 \tfrac{d}{dt} + R_0)w = 0$$

where $R_i \in \mathbb{R}^{p\times q}$ are real matrices. We define

$$R := R_d s^d + ... + R_1 s + R_0.$$

Then R is a polynomial $p \times q$ matrix, and we may rewrite in the concise form

$$R(\tfrac{d}{dt})w = 0.$$

where $\mathbf{R}(s)$ is a matrix of polynomials; $\frac{d}{dt}$ is a differential operator and $w : \mathbb{R} \to \mathbb{R}^q$ is a signals .

Remark 1.2. *We have $\mathbb{R}[s]^{p\times q} = \mathbb{R}^{p\times q}[s]$. To see this we can consider the following 2×3 matrix given below:*

$$R[s] = \begin{bmatrix} s^3 & -2+s & 3 \\ -1+s^2 & 1+s+s^2 & s \end{bmatrix} \in \mathbb{R}[s]^{2\times 3}$$

Then, R can also be written as

$$\begin{aligned} R[s] &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} s^3 + \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} s^2 + \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} s + \begin{bmatrix} 0 & -2 & 3 \\ -1 & 1 & 0 \end{bmatrix} \\ &= R_3 s^3 + R_2 s^2 + R_1 s + R_0. \end{aligned}$$

Where R_i's are the coefficients of s^3 , s^2, s and s^0 respectively. Thus,

$$R(\frac{d}{dt})w = 0 \quad \Leftrightarrow \quad (R_3 \frac{d^3}{dt^3} + R_2 \frac{d^2}{dt^2} + R_1 \frac{d}{dt} + R_0)w = 0$$

The multivariable differential equation $R(\frac{d}{dt})w = 0$ is :

$$\frac{d^3}{dt^3} w_1 - 2w_2 + \frac{d}{dt} w_2 + 3w_3 = 0$$

$$-w_1 + \frac{d^2}{dt^2} w_1 + w_2 + \frac{d}{dt} w_2 + \frac{d^2}{dt^2} w_2 + \frac{d}{dt} w_3 = 0$$

Difference systems: Systems of linear difference equations with constant coefficients. These can be put in the form

$$R_d w(t+d) + ... + R_1 w(t+1) + R_0 w(t) = 0, \forall t \in \mathbf{T}$$

where $R_i \in \mathbb{R}^{p \times q}$ are real matrices. Using the shift operator σ_τ , we may write

$$(R_d \sigma_d + ... + R_1 \sigma_1 + R_0)w = 0.$$

We define $\sigma := \sigma_1$, then $\sigma_k = \sigma^k$ (k-fold application of σ). If we put again

$$R := R_d s^d + ... + R_1 s + R_0$$

then we can write in the concise form

$$R(\sigma)w = 0.$$

Example 1.2.1. *Kepler* (1571−1630) *formulated his laws in a non-differential way. The differential calculus was developed by Newton* (1642 − 1727) *and Leibnitz* (1646 − 1716)*. It was in fact Newton who came up with a differential equation for the motion of planets around the sun (the law of gravity). However, it is quite clear that Kepler's laws describe a non-linear and time-invariant system.*

Example 1.2.2. *The system given by*

$$w(t+1) = Aw(t), \forall t \in \mathbf{T}$$

where

$$A = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

is a linear time-invariant difference system. Using the shift operator, it reads

$$\sigma w = Aw \text{ or } (\sigma I - A)w = 0:$$

The polynomial matrix R takes the form

$$R = sI - A = \begin{bmatrix} s & -1 \\ \frac{-1}{2} & s - \frac{1}{2} \end{bmatrix}$$

1.3 Representations

In the following, let $\mathbf{R}$ be a $p \times q$ polynomial matrix in the variable s, with real coefficients. We write

$$\mathbf{R} \in \mathbb{R}[s]^{p \times q}.$$

In here we will restrict our discussion to the following **standard models**.

- *The continuous standard model* is

$$\mathcal{B} = \{w \in \mathcal{A}^q | \quad \mathbf{R}(\frac{d}{dt})w = 0\}$$

where $\mathcal{A} = \mathcal{D}'(\mathbf{T})$ with $\mathbf{T} = \mathbb{R}$ or $\mathbf{T} = \mathbb{R}_+$.

- *The discrete standard model* is

$$\mathcal{B} = \{w \in \mathcal{A}^q | \quad \mathbf{R}(\sigma)w = 0\}$$

where $\mathcal{A} = \mathbb{R}^{\mathbf{T}}$ with $\mathbf{T} = \mathbb{N}$ or $\mathbf{T} = \mathbb{Z}$.

Now consider a dynamical system $\Sigma = (\mathcal{A}, q, \mathcal{B})$ described by systems of differential equation $\mathbf{R}(\frac{d}{dt})w = 0$. This system is said to be represented/ parameterized by the polynomial matrix $\mathbf{R}(s) \in \mathbb{R}[s]^{p\times q}$.i.e, $\mathbf{R}(s)$ obviously defines Σ. However, there are many polynomial matrices that represent the same dynamical system.

Definition 1.3.1 (Equivalent differential equations)**.** *:*
Let $\mathbf{R}_i(s) \in \mathbb{R}[s]^{p_i\times q}, i = 1, 2$. *Then the differential equations*

$$\mathbf{R}_1(\tfrac{d}{dt})w = 0 \textit{ and } \mathbf{R}_2(\tfrac{d}{dt})w = 0$$

are said to be equivalent if they define the same dynamical system.

Definition 1.3.2 (*Unimodular Matrix*)**.** *A square polynomial matrix* $\mathbf{U}$ *is called unimodular if its determinant is a non-zero constant, that is,* $\det(\mathbf{U}) \in \mathbb{R}\backslash\{0\}$. *This means there exist a polynomial matrix* $\mathbf{V}$ *such that*

$$\mathbf{UV} = \mathbf{VU} = \mathbf{I}.$$

Clearly, $\mathbf{V}$ *is the inverse of* $\mathbf{U}$*, which exists because* $\mathbf{U}$ *is non-singular, i.e.*$\det(\mathbf{U}) \neq 0$.

Example 1.3.1. *Consider the following two matrices:*

$$U = \left[\begin{smallmatrix} s+1 & 1 \\ s & 1 \end{smallmatrix}\right] \qquad \textit{and} \qquad V = \left[\begin{smallmatrix} s+1 & 1 \\ s & s \end{smallmatrix}\right]$$

clearly $\det(U) = s + 1 - s = 1$. *Hence, it is unimodular matrix and its inverse* $\left[\begin{smallmatrix} 1 & -1 \\ -s & s+1 \end{smallmatrix}\right]$ *is also a polynomial matrix.*
But, since $\det(V) = s^2$, *it is not scalar. Hence,* V *is not unimodular matrix. Again its inverse* $\frac{1}{s^2}\left[\begin{smallmatrix} s & -1 \\ -s & s+1 \end{smallmatrix}\right]$ *is not polynomial matrix.*

Question 1: *What makes unimodularity different from non-singularity?*

The answer to this question is obvious that unimodularity is much stronger than non-singularity. The crucial point is that $\mathbf{U}$ possesses a polynomial (rather than a rational) inverse $\mathbf{V}$.

<u>**Question** 2</u>: Why we need to have (learn) unimodular matrices to define system behaviour?

<u>**Answer**</u>: Since many polynomial matrices represent the same dynamical system, once we are given a polynomial matrix $\mathbf{R}$ representing a system behavior, pre-multiplying $\mathbf{R}$ by a unimodular matrix $\mathbf{U}(i.e, \mathbf{UR} = \hat{\mathbf{R}})$ will not affect the system's behavior. Hence, it helps us to find other representations.

$$i.e; \qquad \mathbf{R} \text{ and } \mathbf{UR} = \hat{\mathbf{R}} \text{ represent the same behaviour.}$$
$$\Rightarrow \mathbf{R}(\tfrac{d}{dt})w = 0 \Rightarrow \mathbf{U}(\tfrac{d}{dt})\mathbf{R}(\tfrac{d}{dt})w = \hat{\mathbf{R}}(\tfrac{d}{dt})w = 0$$

and

$$\hat{\mathbf{R}}(\tfrac{d}{dt})w = 0 \Rightarrow \mathbf{V}(\tfrac{d}{dt})\hat{\mathbf{R}}(\tfrac{d}{dt})w = \mathbf{R}(\tfrac{d}{dt})w = 0$$

where $\mathbf{V}$ is the polynomial inverse of $\mathbf{U}$. The same holds if we replace $\frac{d}{dt}$ by σ.
Property of pre-multiplying $R(s)$ by unimodular matrix $U(s)$ has an advantage to bring $R(s)$ into the convenient form,such as triangular form.

Definition 1.3.3 (Minimal representation of $\mathcal{B}$). *A polynomial matrix* $\mathbf{R}$ *is called a minimal representation of* $\mathcal{B}$ *if there exists no polynomial matrix which represents the same behavior and has a smaller number of rows.*

Theorem 1.3.1 (Smith form). *For every polynomial matrix* $\mathbf{R} \in \mathbb{R}[s]^{p\times q}$ *there exist unimodular matrices* $\mathbf{U} \in \mathbb{R}[s]^{p\times p}$ *and* $\mathbf{V} \in \mathbb{R}[s]^{q\times q}$ *such that* $\mathbf{URV} = \left[\begin{smallmatrix} \mathbf{D} & 0 \\ 0 & 0 \end{smallmatrix}\right]$
where $\mathbf{D} \in \mathbb{R}[s]^{r\times r}$ *is a non-singular diagonal matrix*

$$\mathbf{D} = \begin{bmatrix} d_1 & & & \\ & \cdot & & \\ & & \cdot & \\ & & & d_r \end{bmatrix}, with\ d_1|d_2|...|d_r.$$

This notation means that for $i = 1, ..., r-1$, *the polynomial* d_i *divides* d_{i+1}, *that is,* $d_{i+1} = d_i e_i$ *for some polynomial* e_i. *Clearly, the integer* r *is precisely the rank of the matrix* $\mathbf{R}$ *(over the quotient field* $\mathbf{R}(s)$ *of* $\mathbb{R}[s]$*). The matrix* $\left[\begin{smallmatrix} \mathbf{D} & 0 \\ 0 & 0 \end{smallmatrix}\right]$ *is called* ***Smith form*** *of* $\mathbf{R}$.

Proof. (cf:[1]; Appendix D, Smith form) □

Corollary 1.1. *For every polynomial matrix* $\mathbf{R} \in \mathbb{R}[s]^{p\times q}$ *there exists a unimodular matrix* $\mathbf{U} \in \mathbb{R}[s]^{p\times p}$ *such that*

$$\mathbf{UR} = \left[\begin{smallmatrix} R_1 \\ 0 \end{smallmatrix}\right]$$

where $R_1 \in \mathbb{R}[s]^{r\times q}$ *has full row rank, that is, rank*$(R_1) = r$.

Proof. Let $\mathbf{U}$ and $\mathbf{V}$ be unimodular matrices where $\mathbf{U} \in \mathbb{R}[s]^{p\times p}$ and $\mathbf{V} \in \mathbb{R}[s]^{q\times q}$ and set $\mathbf{W} := \mathbf{V}^{-1}$. Then we have

$$\mathbf{URVW} = \left[\begin{smallmatrix} \mathbf{D} & 0 \\ 0 & 0 \end{smallmatrix}\right]\left[\begin{smallmatrix} w_1 \\ w_2 \end{smallmatrix}\right].$$

and we set $R_1 := \mathbf{D}w_1$, which has full row rank, by construction (Since $\mathbf{D}$ is non-singular and w_1 has full row rank). □

Since $\mathbf{R}$ and R_1 represent the same behavior, we have the following conclusion.

Lemma 1.3.1. *Any $\mathcal{B}$ possesses a representation matrix with full row rank. A representation $\mathbf{R}$ of $\mathcal{B}$ is minimal if and only if $\mathbf{R}$ has full row rank.*

Proof. : ($\Rightarrow$:) (*Minimality Implies full row rank*)
Suppose $\mathbf{R}(s)$ is a minimal representation of $\mathcal{B}$.
w.t.s: $\mathbf{R}(s)$ has full row rank.
suppose not ($\mathbf{R}$ has no full row rank)!. This implies that there exist a unimodular matrix $\mathbf{U}(s) \in \mathbb{R}[s]^{p\times p}$ such that

$$\mathbf{UR} = \begin{bmatrix} \bar{\mathbf{R}} \\ 0 \end{bmatrix} \tag{1.1}$$

N.B $\mathcal{B}$ is also represented by (1.1) and hence by $\bar{\mathbf{R}}(\frac{d}{dt})w = 0$. But since the number of rows of $\bar{\mathbf{R}}$ is strictly less than the number of rows of $\mathbf{R}$, $\mathbf{R}$ cannot be minimal. This leads to contradiction to what we supposed. Hence, Minimality implies full row rank.

(:$\Leftarrow$) (**Full row rank implies Minimality**)
Suppose $\mathbf{R}(s)$ has full row rank. We need to show that it is a minimal representation.
Suppose it is not minimal!
$\Rightarrow$ There exist a representation $\bar{\mathbf{R}}(s) \in \mathbb{R}[s]^{\bar{p}\times q}$ with $\bar{p} < p$ where $p = rank(\mathbf{R})$
$\Rightarrow$ There exist a unimodular matrix $\mathbf{U}(s) \in \mathbb{R}[s]^{p\times p}$ such that

$$\mathbf{U}(s)\mathbf{R}(s) = \begin{bmatrix} \bar{\mathbf{R}}(s) \\ 0 \end{bmatrix} \tag{1.2}$$

Note: The row rank of a matrix is invariant under pre-multiplication by unimodular matrix. Hence, the row rank of the right hand side of (1.2) is equal to p. This also implies $\bar{\mathbf{R}}(s)$ should have row rank equal to p which contradicts with what we supposed. Hence, full row rank implies minimality.
Note that the proof of the first part of the lemma immediately follows from corollary 1.1 □

1.4 The fundamental principle

Theorem 1.4.1 (Fundamental principle, square matrix version)**.** *Let $\mathbf{P} \in \mathbb{R}[s]^{p\times p}$ be a non-singular polynomial matrix. Then*

$$\mathbf{P}(\tfrac{d}{dt})y = v \qquad \textit{or} \qquad \mathbf{P}(\sigma)y = v$$

possesses a solution $y \in \mathcal{A}^p$ for every choice of the right hand side signal $v \in \mathcal{A}^p$.

Proof. (cf: [1],page 20) □

Corollary 1.2 (Fundamental principle, non-square matrix version)**.** *Let* $\mathbf{R} \in \mathbb{R}[s]^{p\times q}$ *be a full row rank polynomial matrix, that is, rank* $(\mathbf{R}) = p$*. Then*

$$\mathbf{R}(\tfrac{d}{dt})w = v, \text{ or } \mathbf{R}(\sigma)w = v$$

possesses a solution $w \in \mathcal{A}^q$ *for every choice of the right hand side signal* $v \in \mathcal{A}^p$*.*

Proof. (cf: [1],page 32) □

1.5 Systems with Latent Variables

Mathematical model is a pair $(\mathbb{U}\ ,\ \mathcal{B})$ with $\mathbb{U}$ a set called Universum and elements of it are called outcomes and $\mathcal{B} \subseteq \mathbb{U}$ are called the behaviour. A modeled system may contain two different kinds of variables (components of signal vectors) called Manifest and Latent variables.

Definition 1.5.1. ***(Manifest and Latent variables)***

1. ***Manifest Variables***:*Manifest variables are components of signal vectors in which we are truly interested. These variables describe the system sufficiently*
2. ***Latent Variables***: *They are auxiliary variables introduced during modeling process. They are variables introduced in a system for mathematical convenience, eg. for reducing a system to first order.*

Although introducing latent variables in modelling a system is advantageous for defining system behaviour , it sometimes makes the system costy to handle. A good question to ask is "Why we need to study this topic?" Obviously, once we have a good understanding about a system behaviour and signal vector, the next step will be to analysis components of a signal vector and classify them as essential (manifest) and auxiliary (latent). We now develop a mechanism to make latent variables get rid of our system and have the most precise one only with manifest variables using the fundamental principle.

Definition 1.5.2. *A dynamical system with latent variables is defined as:*

$$\Sigma_L = (\mathbf{T}, \mathbf{W}, \mathbf{L}, \mathcal{B}_f), \text{ Where}$$

T *Time axis*

W *Manifest signal space*

L *Latent variable space*

$\mathcal{B}_f \subseteq (\mathbf{W} \times \mathbf{L})^T$ *full behavior*

It defines a latent variable representation of the manifest dynamical system $\Sigma = (\mathbf{T}, \mathbf{W}, \mathcal{B})$ with the manifest behaviour

$$\mathcal{B} := \{w : \mathbf{T} \to \mathbf{W} | \exists l : \mathbf{T} \to \mathbf{L} \text{ s.t } (w, l) \in \mathcal{B}_f\}$$
$$\mathcal{B}_f \text{ is an internal behaviour}$$
$$\mathcal{B} \text{ is an external behaviour}$$

<u>**N.B**</u>: In the dynamical system with latent variables each trajectory in the full behaviour $\mathcal{B}_f$ consists of a pair (w, l) with $w : \mathbf{T} \to \mathbf{W}$ and $l : \mathbf{T} \to \mathbf{L}$.
Let $w := (w_1 ... w_q)$ and $l := (l_1 ... l_r)$. Then, the joint equations governing w and l is (in here we show only for the continuous time base but still similar argument holds true for the discrete case):

$$\mathbf{R}(\tfrac{d}{dt})w = \mathbf{M}(\tfrac{d}{dt})l$$
$$\text{where } w : \mathbb{R} \to \mathbb{R}^q \text{ denotes the manifest variable and}$$
$$l : \mathbb{R} \to \mathbb{R}^r \text{ denotes the latent variables and}$$
$$\mathbf{R}(s) \in \mathbb{R}[s]^{p \times q} \text{ and } \mathbf{M}(s) \in \mathbb{R}[s]^{p \times r}$$

Ingeneral a behavioural system with latent variable is given by :

$$\mathcal{B} := \{w \in \mathbb{R}^q | \exists l \in \mathbb{R}^r \text{ s.t } \mathbf{R}(\tfrac{d}{dt})w = \mathbf{M}(\tfrac{d}{dt})l\}$$

This also tells us that except their existence, we do not care about the precise form of latent variables.

In here let us ask our self that wether we can write this $\mathcal{B}$ as standard model or not! i.e, can we find a convincing and relevant matrix $\hat{\mathbf{R}}$ such that:

$$\mathcal{B} = \{w \in \mathcal{A}^q | \hat{\mathbf{R}}(\tfrac{d}{dt})w = 0\}$$

Obviously the answer is yes! There is an easy way of obtaining the desired matrix $\hat{\mathbf{R}}$ from the given matrices $\mathbf{R}$ and $\mathbf{M}$. Let us develop some other concepts beforehand so as to generalise.

Consider a vector polynomial $\mathbf{C} \in \mathbb{R}[s]^{1 \times p}$. Then one can solve a linear system of equation $\mathbf{CM} = 0$. This means,one can compute left Kernel of $\mathbf{M}$ over a polynomial ring.

Using the Smith form, we can show that there exist $p - rank(\mathbf{M})$ linearly independent solutions that span this kernel.
Let $c_1 ... c_{p-rank(M)}$ be generating system for the left kernel of $\mathbf{M}$. Collecting the row vectors in matrix $\mathbf{X}$, We have constructed a matrix $\mathbf{X}$ satisfying the following three conditions.

1. $\mathbf{XM} = 0$

2. All $\mathbf{C}$ with $\mathbf{CM} = 0$ can be written as a polynomial Linear combination of rows of $\mathbf{X}$. i.e $\mathbf{C} = \mu\mathbf{X}$ for some polynomial row vector μ.

3. $\mathbf{X}$ has full row rank.

Lemma 1.5.1. *Let $\mathbf{X}_1$ and $\mathbf{X}_2$ be two matrices with the three properties above. Then, we must have $\mathbf{X}_1 = \mathbf{U}\mathbf{X}_2$ for some unimodular matrix $\mathbf{U}$*

Proof. From the first property we have

$$\mathbf{X}_1\mathbf{M} = 0 \text{ and } X_2\mathbf{M} = 0$$

Thus, by condition 2, there exist polynomial matrices $\mathbf{U}$ and $\mathbf{V}$ such that :

$$\mathbf{X}_1 = \mathbf{U}\mathbf{X}_2 \text{ and } \mathbf{X}_2 = \mathbf{V}\mathbf{X_1}$$
$$\Rightarrow \mathbf{X}_1 = \mathbf{U}\mathbf{V}\mathbf{X}_1 \text{ and } \mathbf{X}_2 = VUX_2$$
$$\Rightarrow (\mathbf{I} - \mathbf{U}\mathbf{V})\mathbf{X}_1 = 0 \text{ and } (\mathbf{I} - \mathbf{V}\mathbf{U})\mathbf{X}_2 = 0$$

Since $\mathbf{X}_i$s' have full row rank, we must have that $\mathbf{UV} = \mathbf{VU} = \mathbf{I}$. Hence,$\mathbf{U}$ and $\mathbf{V}$ are unimodular matrices. □

Theorem 1.5.1. *Let $\mathbf{R}$ and $\mathbf{M}$ be given polynomial matrices, with the same number of rows. Let $\mathbf{X}$ be as described above, and define $\hat{\mathbf{R}} := \mathbf{XR}$. Then*

$$\exists l \in \mathcal{A}^r : \mathbf{R}(\tfrac{d}{dt})w = \mathbf{M}(\tfrac{d}{dt})l \Leftrightarrow \hat{\mathbf{R}}(\tfrac{d}{dt})w = 0$$

(analogously if $\frac{d}{dt}$ is replaced by σ)

Proof. By corollary 1.1, there exists a unimodular matrix $\mathbf{V}$ such that

$$\mathbf{VM} = \left[\begin{smallmatrix}\mathbf{M}_1\\ 0\end{smallmatrix}\right]$$

where $\mathbf{M}_1$ has full row rank. Then

$$\exists l \in \mathcal{A}^r : \mathbf{R}(\tfrac{d}{dt})w = \mathbf{M}(\tfrac{d}{dt})l$$
$$\Leftrightarrow \mathbf{V}(\tfrac{d}{dt})\mathbf{R}(\tfrac{d}{dt})w = \mathbf{V}(\tfrac{d}{dt})\mathbf{M}(\tfrac{d}{dt})l$$
$$\Leftrightarrow \left[\begin{smallmatrix}\mathbf{R}_1\\ \mathbf{R}_2\end{smallmatrix}\right](\tfrac{d}{dt})w = \left[\begin{smallmatrix}\mathbf{M}_1\\ 0\end{smallmatrix}\right](\tfrac{d}{dt})l$$
$$\Leftrightarrow \mathbf{R_1}w = \mathbf{M_1}l \text{ and } \mathbf{R_2}w = 0$$

By the fundamental principle for non- square matrix version, since $\mathbf{M}_1$ has full row rank, for any right hand side v (where $v = \mathbf{R}_1 w$), The system

$$M_1 l = v$$

Possesses a solution l. Thus, $\mathbf{R}_1 w = \mathbf{M}_1 l$ is always true. Therefore, we have

$$\left[\begin{smallmatrix}\mathbf{R}_1\\ \mathbf{R}_2\end{smallmatrix}\right](\frac{d}{dt})w = \left[\begin{smallmatrix}\mathbf{M}_1\\ 0\end{smallmatrix}\right](\frac{d}{dt})l \;\;\Leftrightarrow\;\; \mathbf{R}_2(\frac{d}{dt})w = 0$$

It remains to establish a relation between $\mathbf{R}_2$ and $\hat{\mathbf{R}}$. But how?
Define $\hat{\mathbf{X}} := [0, \mathbf{I}]\mathbf{V}$ satisfying the three properties given above. That means there exists a unimodular matrix $\mathbf{U}$ such that

$$\mathbf{X} = \mathbf{U}\hat{\mathbf{X}}$$

Thus

$$\hat{\mathbf{R}} = \mathbf{X}\mathbf{R} = \mathbf{U}\hat{\mathbf{X}}\mathbf{R} = \mathbf{U}[0, I]\mathbf{V}\mathbf{R} = U[0, I]\left[\begin{smallmatrix}\mathbf{R}_1\\ \mathbf{R}_2\end{smallmatrix}\right] = \mathbf{U}\mathbf{R}_2.$$

$$\therefore \hat{\mathbf{R}}(\tfrac{d}{dt})w = 0 \Leftrightarrow \mathbf{R}_2(\tfrac{d}{dt})w = 0$$

This completes the proof. □

1.6 Inputs, Outputs and autonomous Systems

Definition 1.6.1. *Let the signal vector w be partitioned as*

$$w = \left[\begin{smallmatrix}w_1\\ w_2\end{smallmatrix}\right],$$

where $w_i \in \mathcal{A}^{q_i}$, where q_1 and q_2 are positive integers with $q_1 + q_2 = q$. The subvector w_1 is called a vector of free variables (or: inputs) of $\mathcal{B}$ if it is unconstrained by the system law, that is,

$$\forall w_1 \in \mathcal{A}^{q_i}, \exists w_2 \in \mathcal{A}^{q_2} : \left[\begin{smallmatrix}w_1\\ w_2\end{smallmatrix}\right] \in \mathcal{B}.$$

QUESTION: How can we find free variables? and What is the maximal number of free variables of a system $\mathcal{B}$?
Free variables can be found as follows: We may assume that $\mathbf{R} \in \mathbb{R}[s]^{p\times q}$ has full row rank, that is, $rank(\mathbf{R}) = p \leq q$. Then there exists a $p \times p$ non-singular submatrix of $\mathbf{R}$. Assume that $p < q$. We can always permute the columns of $\mathbf{R}$ (this just corresponds to renumbering our signal components) such that

$$R = \left[\begin{smallmatrix}-Q & | P\end{smallmatrix}\right]$$

where $\mathbf{P}$is such a non-singular $p \times p$ matrix. Writing

$$w = \left[\begin{smallmatrix}u\\ y\end{smallmatrix}\right]$$

correspondingly, our system law $\mathbf{R}(\frac{d}{dt})w = 0$ takes the form

$$\mathbf{P}(\tfrac{d}{dt})y = \mathbf{Q}(\tfrac{d}{dt})u.$$

From fundamental principle we obtain that this equation has a solution $y \in \mathcal{A}^p$ for every choice of $u \in \mathcal{A}^m$, where $m := q - p$. Therefore, u is a vector of free variables of $\mathcal{B}$.

When we come to the question of maximal number of free variables of a system $\mathcal{B}$, we need to define the input-dimension of $\mathcal{B}$ prior.

$$\text{idim}(\mathcal{B}) := max\{k \in \mathbb{N} | \exists \Pi \text{ such that } \mathcal{B} \to \mathcal{A}^k, w \mapsto [\mathbf{I}_k, 0]\Pi w \text{ is surjective}\}.$$

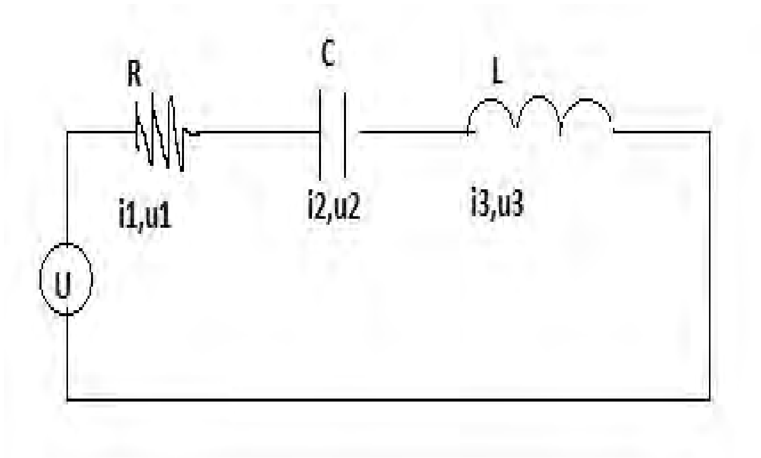
where Π is a permutation matrix. Note that the mapping $\mathcal{B} \to \mathcal{A}^k$ is simply the projection of w onto a subvector consisting of k components. Thus, the input-dimension is the largest integer k such that there exist k components of $w \in \mathcal{B}$ that are free (i.e., unconstrained by the system law). When we speak of the surjectivity of $\mathcal{B} \to \mathcal{A}^k, w \mapsto \mathbf{S}w$, where $\mathbf{S} = [\mathbf{I}_k, 0]\Pi$, we mean that

$$\forall \xi \in A^k, \exists w \in A^q : \left[\begin{smallmatrix} I \\ 0 \end{smallmatrix}\right] \xi = \left[\begin{smallmatrix} S \\ R \end{smallmatrix}\right] (\tfrac{d}{dt}) w.$$

Definition 1.6.2. *A system law in the form of* $\mathbf{P}(\frac{d}{dt})y = \mathbf{Q}(\frac{d}{dt})u.$, *where* $\mathbf{P}$ *is square and non-singular, is called an input-output representation of* $\mathcal{B}$. *One calls* $p = rank(\mathbf{R})$ *the number of outputs (or output-dimension), and* $m = q - p$ *the number of inputs (or input-dimension). The signal subvector* u *is called input,and the signal subvector* y *is called output.*

The input-output representation of a system is not unique. The same behavior may have several different input-output representations. This is due to the fact that there may be more than one way to select a non-singular $p \times p$ submatrix of $\mathbf{R}$, where $p = rank(\mathbf{R})$.

Example 1.6.1. *Consider an electrical circuit given below:*



$$I_{tot} := I = i_1 = i_2 = i_3 and \quad \sum_{i=1}^{3} u_i = U$$

The system behaviour will be written as

$$B = \{(I, U, u_1, u_2, u_3)^T \in \mathcal{A}^5 | u_1 + u_2 + u_3 - U = 0, u_1 = RI, C\dot{u}_2 = I, L\dot{I} = u_3\}$$

The system law can be written being condensed as $\hat{\mathbf{R}}(\frac{d}{dt})w = 0$; *where,*

$$\hat{\mathbf{R}}(\frac{d}{dt})w = \begin{bmatrix} 0 & -1 & 1 & 1 & 1 \\ R & 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & Cs & 0 \\ Ls & 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} I \\ U \\ u_1 \\ u_2 \\ u_3 \end{bmatrix} = 0.$$

This System admits five different input-output representations.This is because $\hat{\mathbf{R}}$ *has rank* 4 *and every* 4×4 *submatrix is non-singular (the parameters* $\boldsymbol{R}$*,* $\boldsymbol{C}$*,* $\boldsymbol{L}$ *are supposed to be positive). Each of the* 5 *signals* I, U, u_1, u_2, u_3 *can take the role of the input, and then the remaining* 4 *signals will be outputs. However, from the physical point of view, one would usually consider* U *as the input and* I, u_1, u_2, u_3 *as outputs.*

Definition 1.6.3 (Autonomous System)**.** *A system which has no free variables is called autonomous system (self governed/ under its own law).*

$$\textit{That is,} \qquad \mathcal{B} \textit{ is autonomous} \Leftrightarrow idim(\mathcal{B}) = 0$$

Lemma 1.6.1. *$\mathcal{B}$ is autonomous if and only if it has a square non-singular representation matrix.*

Proof. If $\mathcal{B}$ is autonomous, then $idim(\mathcal{B}) = 0$, that is, any representation matrix has full column rank. On the other hand, we can always find a representation with full row rank. Combining this, we get that there exists a square representation matrix with full rank, that is, a non-singular matrix. Conversely, let $\mathcal{B}$ be represented by a square non-singular matrix, say, $\mathbf{P} \in \mathbb{R}[s]^{q \times q}$ with $rank(\mathbf{P}) = q$. Then $idim(\mathcal{B}) = q - q = 0$, that is,$\mathcal{B}$ is autonomous. □

Example 1.6.2. *Consider the system given by* $\mathbf{R}(\sigma)w = \begin{bmatrix} \sigma & -1 \\ \frac{-1}{2} & \sigma - \frac{1}{2} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = 0$*. Since* $idim(\mathcal{B}) = p - rank(\mathbf{R}) = 0$*, the system has no free variable and hence it is autonomous.*

Example 1.6.3. *A system given by the scalar equation* $p(\frac{d}{dt})w = 0$*, where* $0 \neq p \in \mathbb{R}[s]$*, is autonomous.*

1.6.1 Solution Technique to Input Output Equation

Let $\mathbf{P}(\frac{d}{dt})y = \mathbf{Q}(\frac{d}{dt})u$ be a system law in input-output form and assume that $u \in \mathbf{A}^m$ is given. From the fundamental principle, we know that there exist outputs belonging to this input. But,what can be said about the set of these outputs?
Let y_p be one particular output that belongs to u, that is, $\mathbf{P}(\frac{d}{dt})y_p = \mathbf{Q}(\frac{d}{dt})u$. Let y be another output that belongs to u, then $\mathbf{P}(\frac{d}{dt})(y - y_p) = 0$, that is, $y_h := y - y_p$ is a solution of the autonomous system defined by $\mathbf{P}(\frac{d}{dt})y_h = 0$. Therefore, any output y that belongs to u can be written in the form $y = y_p + y_h$ where y_p is one particular solution of the inhomogeneous equation $\mathbf{P}(\frac{d}{dt})y = \mathbf{Q}(\frac{d}{dt})u$ and y_h is an arbitrary solution of the corresponding homogeneous equation.

The continuous homogeneous equation

Consider $\mathbf{P}(\frac{d}{dt})y = 0$, where $\mathbf{P} \in \mathbb{R}[s]^{p \times p}$ is non-singular. Let $0 \neq det(\mathbf{P}) \in \mathbf{R}[s]$ be its determinant and let

$$\Lambda = \{\lambda \in \mathbb{C} | det(\mathbf{P})(\lambda) = 0\}$$

denote the set of its zeros. Any solution y of $\mathbf{P}(\frac{d}{dt})y = 0$ has the form

$$y(t) = \sum_{\lambda \in \Lambda} a_\lambda(t) e^{\lambda t}$$

where $a_\lambda \in \mathbb{C}[t]^p$.

The continuous inhomogeneous equation

Suppose that $\mathbf{P}(\frac{d}{dt})h^{(i)} = \mathbf{Q}(\frac{d}{dt})\delta^{(i)}$, where $\delta^{(i)} \in \mathcal{A}^m$ has δ in the i^{th} position and zeros everywhere else. Here, $\delta \in \mathcal{A} = \mathcal{D}'(\mathbf{T})$ denotes the Dirac delta distribution. Define

$$h := [h^{(1)} ... h^{(m)}] \in \mathcal{A}^{p\times m}.$$

Then h is called impulse response (or: fundamental solution), because its columns can be seen as the system's response (output) to an input which is an impulse (delta distribution).

Then a particular output y belonging to the input u is given by the distributional **convolution** (if it exists)

$$y = h * u$$

that is, in the classical case (u, h locally integrable functions, i.e., regular distributions)

$$y(t) = \int_{-\infty}^{\infty} h(t-\tau)u(\tau)d\tau.$$

The existence of the convolution is guaranteed by any of the following conditions:

- If h or u has compact support, the $h * u$ exists.
- Both h and u have their support in $[0, \infty)$. This means that they assign zero to every test function whose support is in $(-\infty, 0)$. If h are u are continuous functions, this means that $h(t) = 0$ and $u(t) = 0$ for all $t < 0$. Then we have

  $$y(t) = \int_0^t h(t-\tau)u(\tau)d\tau,$$

  which is an integral over a compact interval, which exists due to the assumption of continuity. Similarly, the convolution is always well-defined if $\mathbf{T} = \mathbb{R}_+$.

Example 1.6.4. *Consider $\dot{x}(t) = \boldsymbol{A}x(t) + b(t)$. We take b as the input, and x as the output. Then $\boldsymbol{P} = s\mathbf{I} - A$ and $\boldsymbol{Q} = \mathbf{I}$. An impulse response is given by*

$$h(t) = \begin{cases} e^{\mathbf{A}t}, & t \geq 0 \\ 0, & t < 0. \end{cases}$$

for both $\mathbf{T} = \mathbb{R}$ and $\mathbf{T} = \mathbb{R}_+$. If $\mathbf{T} = \mathbb{R}$, assume that u is a continuous function with

$u(t) = 0$ for $t < 0$. Then we have

$$y(t) = (h * u)(t) = \int_0^t e^{A(t-\tau)} u(\tau) d\tau =$$

as a particular output $y = x$ belonging to the input $u = b$.

The discrete homogeneous equation: $(\mathbf{T} = \mathbb{N})$

Consider $\mathbf{P}(\sigma)y = 0$, where $\mathbf{P} \in \mathbb{R}[s]^{p \times p}$ is non-singular. Let $0 \neq det(\mathbf{P}) \in \mathbb{R}[s]$ be its determinant and let

$$\Lambda = \{\lambda \in \mathbb{C} | det(\mathbf{P})(\lambda) = 0\}$$

denote the set of its zeros. Any solution y of $\mathbf{P}(\sigma)y = 0$ has the form

$$y(t) = \sum_{\lambda \in \Lambda} a_\lambda(t) \lambda^t$$

where $a_\lambda \in \mathbb{C}[t]^p$, for all t that are large enough (recall that this restriction is due to the possible presence of 0 in Λ).

The discrete inhomogeneous equation

Suppose that $P(\sigma)h^{(i)} = Q(\sigma)\delta^{(i)}$, where $\delta^{(i)} \in \mathcal{A}^m$ has δ in the i^{th} position and zeros everywhere else. Here, $\delta \in \mathcal{A} = \mathbb{R}^{\mathbf{T}}$ denotes the sequence that takes the value 1 at time zero, and the value zero everywhere else. Define

$$h = [h^{(1)} ... h^{(m)}] \in \mathcal{A}^{p \times m}$$

Then h is called **impulse response** (or: fundamental solution), because its columns can be seen as the system's response (output) to an input which is an impulse (discrete version of delta function). Then a particular output y belonging to the input u is given by the discrete convolution (if it exists)

$$y = h * u$$

that is,

$$y(t) = \sum_{i=-\infty}^{\infty} h(t-i)u(i).$$

We identify $\mathbb{R}^{\mathbb{N}}$ with the set of sequences $\mathbf{a} \in \mathbb{R}^{\mathbb{Z}}$ with $\mathbf{a}(t) = 0$ for all $t < 0$. The existence of the convolution is guaranteed by any of the following conditions:

- At least one of h, u has compact support

- Let $h(t) = 0$ and $u(t) = 0$ for all $t < 0$. Then

$$y(t) = \sum_{i=0}^{t} h(t-i)u(i).$$

is a finite sum. Thus the convolution always exists for $T = \mathbb{N}$.

Example 1.6.5. *Consider $x(t+1) = \mathbf{A}x(t) + b(t)$. We take b as the input, and x as the output. Then $\mathbf{P} = sI - \mathbf{A}$ and $\mathbf{Q} = I$. An impulse response is given by*

$$h(t) = \begin{cases} A^{t-1}, & t \geq 1; \\ 0, & t < 0. \end{cases}$$

If $\mathbf{T} = \mathbb{N}$, or $\mathbf{T} = \mathbb{Z}$ and $u(t) = 0$ for $t < 0$, we have

$$y(t) = (h * u)(t) = \sum_{i=0}^{t} h(t-i)u(i) = \sum_{i=0}^{t} \mathbf{A}^{t-i-1}u(i) = \sum_{i=0}^{t-1} \mathbf{A}^{t-i-1}u(i)$$

as a particular output $y = x$ belonging to the input $u = b$.

Example 1.6.6. *Determine all input-output $(i.o)$ structure of :*

$$\mathcal{B} = \{w \in \mathcal{D}^{'}(\mathbb{R})^3 : \begin{bmatrix} \frac{d}{dt}+1 & \frac{d^2}{dt^2}-1 & 1 \\ \frac{d}{dt} & \frac{d^2}{dt^2}-\frac{d}{dt} & \frac{d}{dt} \end{bmatrix} w = 0\}$$

Solution:

Let $w = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}$. Since $R(\frac{d}{dt}) = \begin{bmatrix} \frac{d}{dt}+1 & \frac{d^2}{dt^2}-1 & 1 \\ \frac{d}{dt} & \frac{d^2}{dt^2}-\frac{d}{dt} & \frac{d}{dt} \end{bmatrix} \in \mathbb{R}[\frac{d}{dt}]^{2\times 3}$ has full row rank, we have three possibilities for the choice of the input u.

1. Let $y = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ and $u = w_3$

   This is not $i.o$ structure because

$$\det \begin{bmatrix} \frac{d}{dt}+1 & \frac{d^2}{dt^2}-1 \\ \frac{d}{dt} & \frac{d^2}{dt^2}-\frac{d}{dt} \end{bmatrix} = 0$$

2. Let $y = \begin{pmatrix} w_1 \\ w_3 \end{pmatrix}$ and $u = w_2$

   This is an $i.o$ structure because

$$\det \begin{bmatrix} \frac{d}{dt}+1 & 1 \\ \frac{d}{dt} & \frac{d}{dt} \end{bmatrix} = \frac{d^2}{dt^2} \neq 0$$

3. Let $y = \begin{pmatrix} w_2 \\ w_3 \end{pmatrix}$ and $u = w_1$

   This is an $i.o$ structure because

$$\det \begin{bmatrix} \frac{d^2}{dt^2}-1 & 1 \\ \frac{d^2}{dt^2}-\frac{d}{dt} & \frac{d}{dt} \end{bmatrix} = \frac{d^3}{dt^3} - \frac{d^2}{dt^2} \neq 0$$

1.7 State and State Representation

1.7.1 State

Central to the state-space notation is the idea of a state. A state of a system is the current value of internal elements of the system, that change separately (but not completely unrelated) to the output of the system. In essence, the state of a system is an explicit account of the values of the internal system components. Here are some examples:

- Consider an electric circuit with both an input and an output terminal. This circuit may contain any number of inductors and capacitors. The state variables may represent the magnetic and electric fields of the inductors and capacitors, respectively.
- In a spring-mass system. The state variables may represent the compression of the spring or the restoring force of the spring.
- In reaction where certain reagents are poured into a mixing container, the output is the amount of the chemical product produced over time. The state variables may represent the amounts of un-reacted chemicals in the container, or other properties such as the quantity of thermal energy in the container (that can serve to facilitate the reaction).

1.7.2 State Variables

When modeling a system using a state-space equation, we first need to define three vectors: Input variables,Output variables and State variables.

1. **Input variables**
   A **SISO** (Single Input Single Output) system will only have a single input value, but a **MIMO** (Multiple Input Multiple Output) system may have multiple inputs. We need to define all the inputs to the system, and we need to arrange them into a vector.
2. **Output Variables**
   This is the system output value, and in the case of MIMO systems, we may have several. Output variables should be independent of one another, and only dependent on a linear combination of the input vector and the state vector.
3. **State variables**
   The state variables represent values from inside the system, that can change over time. In an electric circuit, for instance, the node voltages or the mesh currents can be state variables. In a mechanical system, the forces applied by springs, gravity, and dashpots can be state variables.

We denote the input variables with u, the output variables with y, and the state variables with x. In essence, we have the following relationship:

$$y = f(x, u)$$

Where $f(x, u)$ is our system. Also, the state variables can change with respect to the current state and the system input:

$$\dot{x} = g(x, u)$$

Where $\dot{x}$ is the rate of change of the state variables.

1.7.3 State-Space Equations

In a state-space system representation, we have a system of two equations: an equation for determining the state of the system, and another equation for determining the output of the system. We will use the variable $y(t)$ as the output of the system, $x(t)$ as the state of the system, and $u(t)$ as the input of the system. We use the notation $\dot{x}(t)$ for the first order derivative of the state vector of the system, as dependent on the current state of the system and the current input. Symbolically, we say that there are transforms g and h, that display this relationship:

$$\dot{x}(t) = g[t_0, t, x(t), x(0), u(t)]$$

$$y(t) = h[t, x(t), u(t)]$$

Note:
If $\dot{x}(t)$ and $y(t)$ are not linear combinations of $x(t)$ and $u(t)$, the system is said to be nonlinear.but in here our focus is only on linear system. The first equation shows that the system state change is dependent on the previous system state, the initial state of the system, the time, and the system inputs. The second equation shows that the system output is dependent on the current system state, the system input,and the current time. If the system state change $\dot{x}(t)$ and the system output $y(t)$ are linear combinations of the system state and input vectors, then we can say the systems are linear systems, and we can rewrite them in matrix form:
State Equation

$$\dot{x} = A(t)x(t) + B(t)u(t)$$

Output Equation

$$y(t) = C(t)x(t) + D(t)u(t)$$

since we are dealing with time-invariant system, we can re-write this as follows:

$$\dot{x} = Ax(t) + Bu(t)$$

$$y(t) = Cx(t) + Du(t)$$

This explicit equations are called **state space representations**.The derivation of this equation, state space equation, from $R(\frac{d}{dt})w = 0$ is shown on Appendix. Similar arguments works for discrete time case by using σ notation instead of differential notation. The state equation shows the relationship between the system's current state and its input, and the future state of the system. The Output Equation shows the relationship between the system state and its input, and the output. These equations show that in a given system, the current output is dependent on the current input and the current state.The future state is also dependent on the current state and the current input.
It is important to note, at this point, that the state space equations of a particular system are not unique, and there are an infinite number of ways to represent these equations by manipulating the **A**, **B**, **C** and **D** matrices using row operations. There are a number of "standard forms" for these matrices, however, that make certain computations easier.Converting between these forms will require knowledge of linear algebra.

1.7.4 Description of Coefficient matrix and their Dimension

We have four constant matrices: **A, B, C, and D** in state space representation. We will explain these matrices below:

Matrix A

Matrix **A** is the system matrix, and relates how the current state affects the state change $\dot{x}$. If the state change is not dependent on the current state, **A** will be the zero matrix.

Matrix B

Matrix **B** is the control matrix, and determines how the system input affects the state change. If the state change is not dependent on the system input, then **B** will be the zero matrix.

Matrix C

Matrix **C** is the output matrix, and determines the relationship between the system state and the system output.

Matrix D

Matrix **D** is the feed-forward matrix, and allows for the system input to affect the system output directly. As most of basic feedback system do not have a feed-forward element, we usually make the **D** matrix is the zero matrix.

Matrix Dimensions

Because we are adding and multiplying multiple matrices and vectors together, we need to be absolutely certain that the matrices have compatible dimensions, or else the equations will be undefined. For integer values p, q, and r, the dimensions of the system matrices and vectors are defined as follows: If the matrix and vector dimensions do not

	vectors	Dimension	Matrices	Dimension
1	x	n x 1	A	n x n
2	x'	n x 1	B	n x m
3	y	p x 1	C	p x n
4	u	m x 1	D	p x m

agree with one another, the equations are invalid and the results will be meaningless. Matrices and vectors must have compatible dimensions or they cannot be combined using matrix operations.

1.7.5 Solution to State Space Equation

For differential system: In this case solutions to State Space Equation are heavily rooted in prior knowledge of Ordinary Differential Equations. The state equation is a first-order linear differential equation, or (more precisely) a system of linear differential equations. Because this is a first-order differential equation, we can use results from Ordinary Differential Equations to find a general solution to the equation in terms of the state-variable x. Once the state equation has been solved for x, that solution can be plugged into the output equation. The resulting equation will show the direct relationship between the system input and the system output, without the need to account explicitly for the internal state of the system.

Solving for x(t) With Zero Input

Looking again at the state equation:

$$\dot{x} = \mathbf{A}x(t) + \mathbf{B}u(t)$$

We can see that this equation is a first-order differential equation, except that the variables are vectors, and the coefficients are matrices. However, because of the rules of matrix calculus, these distinctions don't matter. We can ignore the input term (for now), and rewrite this equation in the following form:

$$\frac{dx(t)}{dt} = \mathbf{A}x(t)$$

And we can separate out the variables as such:

$$\frac{dx(t)}{x(t)} = \mathbf{A}dt$$

Integrating both sides we get the final solution to this equation as:

$$x(t) = e^{\mathbf{A}(t-t_0)}x(t_0)$$

Solving for x(t) With Non-Zero Input

If, however, our input is non-zero (as is generally the case with any interesting system), our solution is a little bit more complicated. Notice that now we have our input term in the equation, we will no longer be able to separate the variables and integrate both sides easily.

$$\dot{x}(t) = \ \mathbf{A}x(t) + \mathbf{B}u(t)$$

We subtract to get the $\mathbf{A}$x(t) on the left side, and then we do something curious; we pre-multiply both sides by the inverse state transition matrix ($e^{-\mathbf{A}t}$):

$$e^{-\mathbf{A}t}\dot{x}(t) - e^{-\mathbf{A}t}\mathbf{A}x(t) = e^{-\mathbf{A}t}\mathbf{B}u(t)$$

The rationale for this last step may seem fuzzy at best, so we will illustrate the point with an example:

Example 1.7.1. *Take the derivative of the following with respect to time:*

$$\frac{d}{dt}(e^{-\boldsymbol{A}t}x(t)) = e^{-\boldsymbol{A}t}\dot{x}(t) - e^{-\boldsymbol{A}t}\boldsymbol{A}x(t)$$

If we look at this result, it is the same as the left hand side of our equation above. Using the result from our example, we can condense the left side of our equation into a derivative:

$$\frac{d(e^{-\mathbf{A}t}x(t))}{dt} = e^{-\mathbf{A}t}\mathbf{B}u(t)$$

Now we can integrate both sides, from the initial time (t_0) to the current time (t), using a dummy variable τ . Finally, if we pre-multiply by $e^{\mathbf{A}t}$, we get our final result:
General State Equation Solution

$$x(t) = e^{\mathbf{A}(t-t_0)}x(t_0) + \int_{t_0}^{t} e^{\mathbf{A}(t-\tau)}\mathbf{B}u(\tau)d\tau$$

If we plug this solution into the output equation, we get:
General Output Equation Solution

$$y(t) = \mathbf{C}e^{\mathbf{A}(t-t_0)}x(t_0) + \mathbf{C}\int_{t0}^{t}(e^{\mathbf{A}(t-\tau)}\mathbf{B}u(\tau)d\tau) + \mathbf{D}u(t)$$

This is the general Time-Invariant solution to the state space equations, with non-zero input.

Difference System: In here we can find a general time-invariant solution for the discrete time difference equations. Let us start working up a pattern. We know the discrete state equation:

$$x(t+1) = Ax(t) + Bu(t)$$

Starting from time $t = 0$, we can start to create a pattern:

$$x(1) = Ax(0) + Bu(0)$$

$$x(2) = Ax(1) + Bu(1) = A^2x(0) + ABu(0) + Bu(1)$$

$$x(3) = Ax(2) + Bu(2) = A^3x(0) + A^2Bu(0) + ABu(1) + Bu(2)$$

With a little algebraic trickery, we can reduce this pattern to a single equation:
General State Equation Solution

$$x(t) = A^tx(0) + \sum_{i=0}^{t-1} A^{t-1-i}Bu(i)$$

Substituting this result into the output equation gives us:
General Output Equation Solution

$$y(t) = CA^tx(0) + \sum_{i=0}^{t-1} CA^{t-1-i}Bu(i) + Du(t)$$

Chapter 2

OBSERVABILITY

What you observe is not only
what you don't know and what you
know is not only what you observed.
My Mother Yesriba Jundi (1959)

Observability, together with other system properties, represent a major concepts of modern control system theory. Even though its concepts is almost abstractly defined, we now intuitively understand its meaning and produce some mathematical check up tests to define observability more rigorously. Our intention is to reduce mathematical derivations and the number of definitions, but at the same time to derive and define both of them very clearly. In that respect, we start with some necessary concepts of observability derivations for linear discrete-time invariant systems and give the corresponding definition. The observability of linear discrete systems is very naturally introduced using only elementary linear algebra. This approach will be extended to continuous-time system observability, where the derivatives of measurements (observations) have to be used. This relaxation concept was introduced by Hungarian control scientist R. E. Kalman in 1960. Here are some definitions I found on internet

- The linear system $\dot{x} = Ax + Bu$, $y = Cx + Du$ is observable if, for any time $\tau > 0$, it is possible to uniquely determine $x(t)$ for $t \in [0, \tau]$ based on knowledge of $u(t)$ and $y(t)$ for $t \in [0, \tau]$.
- A linear time invariant dynamic system (LTIDS)is observable if given $u(t)$, $y(t)$, $t \in [0, \tau]$, $0 \leq \tau < \infty$, we can uniquely determine any initial state $x(0)$ from the input and output history over a finite time interval τ.
- In order to see what is going on inside the system under observation,the system must be observable.
- A system is completely state-observable at time t_0 or the pair (A, C) is observable at t_0 if the only state that is unobservable at t_0 is the zero state $x = 0$.

The basic characteristics of a dynamical system is that they have a memory. The future behaviour of a system cannot be described only by just giving future inputs but along

with it you need all the past inputs starting from the time that the system is constructed. When dealing with a dynamical system, we need to have some essential information regarding the past before we predict the future under the influence of certain excitation. That is why we need to study the topic,Observability.
In order to have a good start on the topic let us have some prelude prior.

2.1 Basic Notions for State Space Systems

Consider the following state space equations

$$\begin{aligned} \dot{x} &= Ax + Bu \qquad \text{or} \qquad & \sigma x &= Ax + Bu \\ y &= Cx + Du & y &= Cx + Du \end{aligned}$$

where $A \in \mathbb{R}^{n\times n}$,$B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$ and $D \in \mathbb{R}^{p\times m}$. Let $X = \mathbb{R}^n$ and $Y = \mathbb{R}^p$

Now we define the following two maps

The state transition map

$$\varphi : \{(t, t_o) \in \mathrm{T}^2 |\ t \geq t_0\} \times X \times U \to X,\ (t, t_0, x_0, u) \mapsto \varphi(t, t_0, x_0, u)$$

This map gives the state at time t if the state at t_0 was x_0 and the control function u was applied.

The state-to-out map

$$\eta : \{(t, t_o) \in \mathrm{T}^2 |\ t \geq t_0\} \times X \times U \to Y,\ (t, t_0, x_0, u) \mapsto \eta(t, t_0, x_0, u)$$

This map gives the output at time t if the state at t_0 was x_0 and the control function u was applied.

Having this two mappings in mind, let us go back to state space system.

$$x(t) = \varphi(t, t_0, x_0, u) = e^{A(t-t_0)}x(t_0) + \int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)d\tau$$

$$OR$$

$$x(t) = \varphi(t, t_0, x_0, u) = A^{(t-t_0)}x(t_0) + \sum_{i=t_0}^{t-1} A^{(t-i-1)}Bu(i)$$

$$y = Cx + Du \qquad \Rightarrow \qquad \eta(t, t_0, x_0, u) = C\varphi(t, t_0, x_0, u) + Du(t).$$

$$\Rightarrow \eta(t, t_0, x_0, u) = Ce^{A(t-t_0)}x(t_0) + \int_{t_0}^{t} Ce^{A(t-\tau)}Bu(\tau)d\tau + Du(t).$$

And for the discrete case we have:

$$\eta(t, t_0, x_0, u) = CA^{(t-t_0)}x(t_0) + \sum_{i=t_0}^{t-1} CA^{(t-i-1)}Bu(i) + Du(t).$$

NB: The state-to-output map is causal. That means if two input functions, u_1 and u_2 coincide(agree) for all $t \in [t_0, t_1]$, then

$$\eta(t_1, t_0, x, u_1) = \eta(t_1, t_0, x, u_2) \text{ for all } x.$$

And also we say that it is strictly causal if the above condition is implied by $u_1(t) = u_2(t)$ for all $t \in [t_0, t_1)$.

Remark 2.1. *In state space system strict causality holds if and only if $D = 0$.*
Justification:
Note that $\eta(t, t_0, x_0, u) \in \mathbf{Y}$; where $\mathbf{Y}$ is an output space. And hence, we have the following.

$$\eta(t_1, t_0, x, u_1) = Ce^{A(t_1-t_0)}x(t_0) + \int_{t_0}^{t_1} Ce^{A(t_1-t)}Bu_1(t)dt + Du_1(t_1).$$

And also

$$\eta(t_1, t_0, x, u_2) = Ce^{A(t_1-t_0)}x(t_0) + \int_{t_0}^{t_1} Ce^{A(t_1-t)}Bu_2(t)dt + Du_2(t_1)$$

Since

$$u_1(t) = u_2(t) \textit{ for all } t \in [t_0, t_1)$$

and

$$\textstyle\int_{t_0}^{t_1} Ce^{A(t_1-t)}Bu_1(t)dt \text{ - } \int_{t_0}^{t_1} Ce^{A(t_1-t)}Bu_2(t)dt = 0,$$

subtracting one from another Yields:

$$\eta(t_1, t_0, x, u_1) - \eta(t_1, t_0, x, u_2) = Du_1(t_1) - Du_2(t_1) = D(u_1(t_1) - u_2(t_1))$$
$$\eta(t_1, t_0, x, u_1) - \eta(t_1, t_0, x, u_2) = 0 \textit{ if and only if } D = 0.$$

Hence, the state-to-output map is strictly causal if $u_1(t) = u_2(t)$ for all $t \in [t_0, t_1)$ implies

$$\eta(t_1, t_0, x, u_1) = \eta(t_1, t_0, x, u_2) \quad \forall x(t).$$

The other and the most important thing that we should know about the system that we are dealing with is its *linearity*. That means:

$$\eta(t, t_0, \lambda_1 x_1 + \lambda_2 x_2, \lambda_1 u_1 + \lambda_2 u_2) = \lambda_1 \eta(t, t_0, x_1, u_1) + \lambda_2 \eta(t, t_0, x_2, u_2), \forall t \geq t_0, \lambda_i \in \mathbb{R}, u_i \in U.$$

Having what we have discussed so far in mind, let us ask ourself " *What kind of problems is observability concerned with? and what does it mean by observability of a system?*"

In order to facilitate this central idea of our discussion, OBSERVABILITY , let us get things down and have some prelude which may help us understand the concept.
Suppose we have a behavior of which the variables are partitioned as $w = \binom{w_1}{w_2}$ .We can say that w_2 is observable from w_1 if w_1, together with the law of the system, determines w_2 uniquely. That means that for each w_1 there exist at most one w_2 such that $\binom{w_1}{w_2}$ belongs to the behaviour. A direct implication of w_2 being observable from w_1 is that in fact all the information of a trajectory w is already contained in the first component w_1.
In our previous chapters, we have seen system with latent variables:

$$B_l = \{w \in A^q | \exists l \in A^r \text{ s.t } R(\tfrac{d}{dt})w = M(\tfrac{d}{dt})l\}$$

and state space system.

- In state space system, the manifest variables are the input and the output. The input is free and can be chosen by the control engineer, and the output is the systems response which can be measured.
- In state space system, the latent variables are the state. They are not directly measurable. We know that they are auxiliary variables introduced during mathematical modeling or simply for mathematical convenience, for instance for reducing a system to first order. In general, the physical meaning of latent variables is vague.

We usually know the systems manifest variables. Hence, It is normal to ask ourself the following question:

What can we conclude about the latent variables from our knowledge of manifest variables?

An observable system is one in which the latent variables can be reconstructed from the manifest variables. Due to the property of the state,we only need to reconstruct the state at specific time t_0, then we know it every where in the future. i.e $\forall t \geq t_0$.
How do we reconstruct the initial state $x(0)$? To answer this question we need the following concepts.

2.2 Distinguishability of states and Observability of a system

Definition 2.2.1. *Let $t_0 \in \mathbf{T}$ be fixed.We can say that the state $x \in \mathbf{X}$ is distinguishable from state $\acute{x} \in \mathbf{X}$ in time $\tau \in \mathbf{T}(\tau \geq 0)$ if there exist $u \in \mathcal{U}$ and $t \in [t_0, t_0 + \tau]$ such that:*

$$\eta(t, t_0, x, u) \neq \eta(t, t_0, \acute{x}, u)$$

In this case we say that u distinguishes between x and $\acute{x}$ in time τ.
NB: Distinguishability of x from $\acute{x}$ is guaranteed if the above condition holds for at least one $\tau \geq 0$.
We say that the system is observable if for any $x \neq \acute{x} \in \mathbf{X}$, the state x can be distinguished from the state $\acute{x}$

Remark 2.2. *Since we are dealing with linear time invariant dynamic system, the starting time t_0 is not important and hence we can put $t_0 = 0$.*

Now let us define the set $\mathcal{J}$ where

$$\mathcal{J}(\tau, x) := \{\acute{x} \in \mathbf{X} | \acute{x} \text{ is indistinguishable from } x \text{ in time } \tau\}$$

and from this we can easily see that $\mathcal{J}(\tau) := \mathcal{J}(\tau, 0)$ denotes the set of states that are indistinguishable from state 0 in time τ. Then we have:

$$x \in \mathcal{J}(\tau) \Leftrightarrow \eta(t, 0, x, u) = \eta(t, 0, 0, u), \forall t \in [0, \tau] \text{ and } u \in \mathbf{U}.$$

Because of the linearity of the system, we can also write

$$\begin{aligned} \eta(t, 0, x, u) &= \eta(t, 0, x, 0) + \eta(t, 0, 0, u) \\ &= \eta(t, 0, x, 0) + \eta(t, 0, x, u) \end{aligned}$$

$$\Rightarrow \eta(t, 0, x, 0) = 0.$$

Therefore,

$$x \in \mathcal{J}(\tau) \Leftrightarrow \eta(t, 0, x, 0) = 0 \quad \forall t \in [0, \tau]. \tag{2.1}$$

Finally,

$$\mathcal{J} := \bigcap_{\tau \geq 0} \mathcal{J}(\tau)$$

is the set of all states that are indistinguishable from state 0.
Hence,

$$x \in \mathcal{J} \Leftrightarrow \eta(t, 0, x, 0) = 0, \forall t \geq 0.$$

Again since,by definition $\eta(t, 0, 0, 0) = 0, \forall t \geq 0$, combining this with the above argument we have:

$$x \in \mathcal{J} \Leftrightarrow \eta(t, 0, x, 0) = \eta(t, 0, 0, 0), \forall t \geq 0.$$

This means that x cannot be distinguished from 0 if and only if the zero input function does not distinguish between x and 0. In other words, If x can be distinguished from 0 at all,then it can also be distinguished from 0 by the zero input function. This can be shown as follow.
Suppose x can be distinguished from 0. Then $\exists u$ and $\exists t' \geq 0$ such that

$$\begin{aligned} &\eta(t', 0, x, u) \neq \eta(t', 0, 0, u) \\ &\qquad \Rightarrow \qquad \eta(t', 0, x, 0) + \eta(t', 0, 0, u) \neq \eta(t', 0, 0, u) \\ &\qquad \Rightarrow \qquad \eta(t', 0, x, 0) \neq 0 \end{aligned}$$

Since $\eta(t', 0, 0, 0) = 0 \ \forall t \geq 0$, we get

$$\eta(t', 0, x, 0) \neq \eta(t', 0, 0, 0)$$

$\Rightarrow$ The zero input $u \equiv 0$ distinguishes between x and 0.

Hence, this implies that the choice of the input function plays no role for the question of observability, *i.e*, we may put $u = 0$, without loss of generality, in our discussion.

Theorem 2.2.1. *Let $s, t \in \mathbf{T}, 0 \leq s \leq t$. Then we have:*

1. $\mathcal{J}(t) \subseteq \mathcal{J}(s)$.
2. $\mathcal{J}(t), \mathcal{J}$ *are subspaces of* $\mathbf{X} = \mathbb{R}^n$.
3. $\exists \tau^* \in \mathbf{T}, \tau^* \geq 0$ *such that* $\mathcal{J} = \mathcal{J}(\tau), \forall \tau \geq \tau^*$.

Proof. 1. Let $x \in \mathcal{J}(t) \Rightarrow x$ is indistinguishable from 0 in time t.

$$\Rightarrow \eta(t^*, 0, x, u) = \eta(t^*, 0, 0, u), \forall t^* \in [0, t], \forall u \in \mathbf{U}.$$

Since $s \in [0, t]$, this again implies $\eta(t^*, 0, x, u) = \eta(t^*, 0, 0, u)$ $\forall t^* \in [0, s], \forall u \in \mathbf{U}$. Hence

$$x \in \mathcal{J}(s).$$

Therefore $x \in \mathcal{J}(t) \Rightarrow x \in \mathcal{J}(s)$, which means $\mathcal{J}(t) \subseteq \mathcal{J}(s)$.

2. Clearly, $\mathcal{J}(t)$ is non-empty, because $0 \in \mathcal{J}(t)$. Let $x_1, x_2 \in \mathcal{J}(t)$ and $\lambda_1, \lambda_2 \in \mathbb{R}$. Then, by equation (2.1) we have

$$\eta(t^*, 0, x_1, 0) = 0 = \eta(t^*, 0, x_2, 0), \ \forall t^* \in [0, t],$$

Now, for $t^* \in [0, t]$, we have

$$\eta(t^*, 0, \lambda_1 x_1 + \lambda_2 x_2, 0) = \lambda_1 \eta(t^*, 0, x_1, 0) + \lambda_2 \eta(t^*, 0, x_2, 0) = 0$$

$$\Rightarrow \lambda_1 x_1 + \lambda_2 x_2 \in \mathcal{J}(\tau).$$

Hence, $\mathcal{J}(\tau)$ is a subspace of $\mathbf{X}$. Similarly, since intersection of subspaces is a subspace, $\mathcal{J}$ is also a subspace of $\mathbf{X}$.

3. Consider a strictly increasing sequence

$$0 = \tau_0 < \tau_1 < \tau_2 < ....$$

in $\mathbf{T}$ with $\lim_{i \to \infty} \tau_i = \infty$.
By part 1,

$$\mathcal{J}(\tau_0) \supseteq \mathcal{J}(\tau_1) \supseteq \mathcal{J}(\tau_2) \supseteq ...$$

By part 2, this is a sequence of subspaces of $\mathbf{X} = \mathbb{R}^n$, with

$$n \geq dim\mathcal{J}(\tau_0) \geq dim\mathcal{J}(\tau_1) \geq dim\mathcal{J}(\tau_2)...$$

This is a decreasing sequence of integers less than or equal to n and greater than or equal to 0. Such a sequence must become stationary, that is, there exists i_0 such that

$$dim\mathcal{J}(\tau_i) = dim\mathcal{J}(\tau_{i_0}) \text{ for all } i \geq i_0.$$

We use the following fact from linear algebra: If a vector space is contained in another vector space of the same finite dimension, then the two vector spaces must be the same. Thus

$$\mathcal{J}(\tau_i) = \mathcal{J}(\tau_{i_0}) \text{ for all } i \geq i_0.$$

For any $\tau \in T, \tau \geq \tau_{i_0}$, there exists $j \geq i_0$ such that $\tau \leq \tau_j$. Then

$$\mathcal{J}(\tau_{i_0}) = \mathcal{J}(\tau_j) \subseteq \mathcal{J}(\tau) \subseteq \mathcal{J}(\tau_{i_0}).$$

We conclude that

$$\mathcal{J}(\tau) = \mathcal{J}(\tau_{i_0}) \text{ for all } \tau \geq \tau_{i_0}$$

and thus

$$\mathcal{J} = \bigcap_{\tau \geq 0} \mathcal{J}(\tau) = \bigcap_{\tau \geq \tau_{i_0}} \mathcal{J}(\tau) = \mathcal{J}(\tau_{i_0}).$$

Put $\tau^* := \tau_{i_0}$, then for $\tau \geq \tau^*$, we have

$$\mathcal{J} \subseteq \mathcal{J}(\tau) \subseteq \mathcal{J}(\tau^*) = \mathcal{J}$$

and thus $\mathcal{J}(\tau) = \mathcal{J}$ for all $\tau \geq \tau^*$

□

Corollary 2.1. *In discrete time,*

$$\mathcal{J}(n-1) = \mathcal{J}$$

where n is the dimension of the state space. In continuous time,

$$\mathcal{J}(\epsilon) = \mathcal{J}$$

for every $\epsilon > 0$.

Corollary 2.2. *The following are equivalent.*

1. *The system is observable*
2. *Any non-zero state can be distinguished from 0,that is $\mathcal{J} = \{0\}$*

Proof. $(1 \Rightarrow 2)$. We are given that the system is observable.

$$\Rightarrow \forall x_1 \neq x_2 \in \mathbf{X}, x_1 \text{ can be distinguished from } x_2 \text{ in time } 0 \leq \tau \in \mathbf{T}$$
$$\Rightarrow \exists u \in \mathbf{U} \text{ and } t \in [0, \tau], \text{ such that } \eta(t, 0, x_1, u) \neq \eta(t, 0, x_2, u)$$

Taking them to one side and applying the linear property of the system we will have

$$\eta(t, 0, x_1 - x_2, 0) \neq 0.$$
$$\Rightarrow x_1 - x_2 \neq 0 \text{ can be distinguished from state } 0.$$
$$\Rightarrow \text{Any non zero state can be distinguished from Zero. That means } \mathcal{J}\{0\}.$$

$(2 \Rightarrow 1)$. Suppose every non zero state can be distinguished from zero.

WTS: The system is observable.

Let x_1 and x_2 be in $\mathbf{X}$ with $x_1 \neq x_2$. Then, $x_1 - x_2 \neq 0$. By assumption, $\exists u$ and $t \in [0, \tau]$, where $\tau = \epsilon > 0$ for continuous case and $\tau = (n-1)$ for discrete case, such that

$$\eta(t, 0, x_1 - x_2, u) \neq \eta(t, 0, 0, u)$$

$$\Rightarrow \eta(t, 0, x_1, u) - \eta(t, 0, x_2, 0) \neq \eta(t, 0, 0, u)$$

$$\Rightarrow \eta(t, 0, x_1, u) \neq \eta(t, 0, x_2, 0) + \eta(t, 0, 0, u)$$

$$\Rightarrow \eta(t, 0, x_1, u) \neq \eta(t, 0, x_2, u)$$

Therefore, x_1 is distinguishable from x_2 and thus the system is observable.

□

2.3 Observability Gramians and Kalmans Observability Matrix

The concept of observability is related to linear systems of algebraic equation. It is well known that a solvable system of linear algebraic equation has a solution if and only if the rank of the system matrix is full. Observability test will also be connected to the rank test of certain matrix,Observability Matrix.

2.3.1 Kalmans Observability Matrix

Consider the following input free linear time invariant system in the state space form.

$$\dot{x}(t) = Ax(t) \qquad or \qquad x(t+1) = Ax(t) \tag{2.2}$$
$$y(t) = Cx(t) \qquad\qquad y(t) = Cx(t) \tag{2.3}$$

where in both cases the initial state $x(0) = x_0 =$ Unknown.

In here, the natural question to be asked is: *Can we learn everything about the dynamical behaviour of the state space variables defined in the above equation by using only*

information from output measurements, $y(t) = Cx(t)$?
clearly, if we know $x(0)$, then the recursion formula $\dot{x}(t) = Ax(t)$ or $x(t+1) = Ax(t)$ apparently provides us complete knowledge about the state variables at any time instant. Thus,the only thing that we have to determine in state measurements is the initial state vector $x(0)$.
Since the n-dimensional vector x_0 has n-unknown components, it is expected that n-measurements are sufficient to determine x_0.

For Discrete Systems

Take $t = 0, 1, 2, ...(n-1)$ and generate the following sequence.

$$
\begin{aligned}
&y(0) = Cx(0)\\
&y(1) = Cx(1) = CAx(0)\\
&y(2) = Cx(2) = CAx(1) = CA^2x(0)\\
&.\\
&.\\
&.\\
&y(n-1) = CA^{(n-1)}x(0)
\end{aligned}
$$

Writing this in matrix form we will have:

$$
\begin{bmatrix} y(0) \\ y(1) \\ y(2) \\ . \\ . \\ . \\ y(n-1) \end{bmatrix}^{np\times 1} = \begin{bmatrix} C \\ CA \\ CA^2 \\ . \\ . \\ . \\ CA^{(n-1)} \end{bmatrix}^{np\times n} x(0)
$$

As we have stated before, our knowledge of linear algebra tells us that system of linear algebraic equations with n unknowns has a unique solution if and only if the system matrix has rank n. Hence, for the above equation to have a unique solution the coefficient matrix of $x(0)$ must be of full column rank . That means,

$$
\text{Rank} \begin{bmatrix} C \\ CA \\ CA^2 \\ . \\ . \\ . \\ CA^{(n-1)} \end{bmatrix} = n
$$

The initial condition $x(0)$ is uniquely determined if the so called Observability matrix

$$
\mathcal{O}(A, C) = \begin{bmatrix} C \\ CA \\ CA^2 \\ . \\ . \\ . \\ CA^{(n-1)} \end{bmatrix}^{np\times n}
$$

has rank n, that is $\text{Rank}(\mathcal{O}) = n$

From this we can see that observability of a system depends on the systems matrix A and the output matrix C. Recall that a matrix and its transpose have the same rank. Hence, we have also the following convenient way of writing observability matrix.

$$
\mathcal{O}(A, C) = \begin{bmatrix} C^T & A^TC^T & (A^2)^TC^T & \dots & (A^{(n-1)})^TC^T \end{bmatrix}
$$

Example 2.3.1. *Consider the following system with the measurement*

$$\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$
$$y(t) = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

Check that wether this system is observable or not.

Solution

The Observability matrix for this system is:

$$\mathcal{O} = \begin{bmatrix} C^T & A^T C^T \end{bmatrix} = \begin{bmatrix} 1 & 7 \\ 2 & 10 \end{bmatrix}$$

From the right hand side matrix we can easily see that the rank($\mathcal{O}$) $= 2 = n$. we can also easily see this by one of the following methods.

- Rows of the matrix $\mathcal{O}$ are linearly independent.
- $det(\mathcal{O}) = -4 \neq 0 \Leftrightarrow \mathcal{O}$ has full $rank = n$. Note that this method is applicable only if the observability matrix is square matrix.

Therefore, the system under consideration is observable.

For Continuous Systems

Following the same arguments as in the case of discrete-time system, we can conclude that the knowledge of x_0 is sufficient to determine $x(t)$ at any time instant, since the recursion $x(t) = e^{A(t-t_0)}x_0$ gives a complete knowledge about the state variable ones initial state is known.

The problem that we face with is to find x_0 from the available measurements $y(t)$. We have solved this problem for discrete-time systems by generating the sequence of measurements at discrete-time instants $t = 0, 1, 2, ...n-1$. Note that a time shift in the discrete-time corresponds to a derivative in the continuous time.An analogous technique in the continuous-time domain is obtained by taking derivatives of the continuous-time measurements.

$$\begin{aligned}
&y(t_0) = Cx(t_0) \\
&\dot{y}(t_0) = C\dot{x}(t_0) = CAx(t_0) \\
&\ddot{y}(t_0) = C\ddot{x}(t_0) = CA\dot{x}(t_0) = CA^2x(t_0) \\
&. \\
&. \\
&. \\
&y^{(n-1)}(t_0) = CA^{n-1}x(t_0)
\end{aligned}$$

This can be put in the matrix form as

$$\begin{bmatrix} y(t_0) \\ \dot{y}(t_0) \\ \ddot{y}(t_0) \\ \vdots \\ y^{(n-1)}(t_0) \end{bmatrix}^{np\times 1} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}^{np\times n} x(t_0) = \mathcal{O}x(t_0)$$

The initial condition $x(t_0)$ can be determined uniquely if and only if the observability matrix has full column rank. That means;

$$\text{Rank}\begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{(n-1)} \end{bmatrix} = n$$

Example 2.3.2. *Given the following state variables of equations of a system. Justify whether the system is observable or not.*

$$\begin{aligned} \dot{x}_1 &= -3x_1 - x_2 + u \\ \dot{x}_2 &= 2x_1 \\ y &= x_1 + u \end{aligned}$$

Justification

First let us write the equation in standard state space equation form.

$$\begin{aligned} \dot{x} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} &= \begin{bmatrix} -3 & -1 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u \\ y &= \begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \end{bmatrix} u \end{aligned}$$

Now the observability matrix for this system is: $\mathcal{O} = \begin{bmatrix} C^T & A^T C^T \end{bmatrix} = \begin{bmatrix} 1 & -3 \\ 0 & -1 \end{bmatrix}$
observe that $Rank(\mathcal{O}) = 2$. Hence the system is observable.

Example 2.3.3 (*Observation of population*)**.** *Consider two populations whose sizes are changing according to the equations*

$$\dot{x}_1 = \lambda_1 x_1, \qquad \dot{x}_2 = \lambda_2 x_2.$$

And suppose we observe $x_1 + x_2$, What can you conclude about the observability of individual population?
In here we have

$$A = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 1 \end{bmatrix} \qquad \text{and} \qquad \mathcal{O}(A, C) = \begin{bmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{bmatrix}$$

and so the individual populations are observable if $\lambda_1 \neq \lambda_2$.

2.3.2 Observability Gramians

The observability Gramian is also another method used to determine whether or not a linear system is observable.
We define the observability Gramians $W(t) \in \mathbb{R}^{n\times n}$ as:

$W(t) = \int_0^t e^{A^T\tau}C^TCe^{A\tau}d\tau$ or $W(t) = \sum_{i=0}^t (A^T)^iC^TCA^i$

Lemma 2.3.1. *$W(\tau)$ is positive definite if and only if there is no vector $v \neq 0$ such that:*

$$Ce^{At}v \equiv 0,\ \forall t \in [0,\tau]$$

Proof. ($\Rightarrow$:) Given that $W(\tau)$ is positive definite. We need to show that there exist no vector $v \neq 0$ such that $Ce^{At}v \equiv 0$, $\forall t \in [0,\tau]$.
Suppose there is such a vector $v \neq 0$ such that $Ce^{At}v \equiv 0$, $\forall t \in [0,\tau]$. Then,by positive definiteness of the Gramian we can write

$$v^TW(t)v = \int_0^\tau v^Te^{A^Tt}C^TCe^{At}vdt > 0$$

$$\Rightarrow \int_0^\tau (Ce^{At}v)^T(Ce^{At}v)dt > 0$$

This leads to contradiction as the integrand is zero and hence, the integral cannot be greater than zero. Hence there exist no vector $v \neq 0$ such that $Ce^{At}v \equiv 0$, $\forall t \in [0,\tau]$.
($\Leftarrow$) Given that there exist no vector $v \neq 0$ such that $Ce^{At}v \equiv 0$, $\forall t \in [0,\tau]$. We need to show that $W(\tau)$ is positive definite.
Since for every vector different from zero the given condition cannot be equal to zero (i.e $Ce^{At}v \neq 0$,). Hence, $\forall v \neq 0$

$$v^TW(t)v = \int_0^\tau v^Te^{A^Tt}C^TCe^{At}vdt > 0$$

Hence,$W(\tau)$ is positive definite. □

Theorem 2.3.1. *The pair (A,C) is observable on $[0,t_f]$ if and only if the observability Gramian $W(t_f)$ is positive definite.*

Proof. (proof of Sufficiency:) Suppose $W(t_f)$ is positive definite and let x_0 be the initial state. Then from output condition (equation 2.3) we have

$$y(t) = Cx(t) = Ce^{At}x_0$$

Multiplying both sides from left by $e^{A^Tt}C^T$, we get

$$e^{A^Tt}C^Ty(t) = e^{A^Tt}C^TCe^{At}x_0$$

integrating from 0 to t_f

$$\int_0^{t_f} e^{A^Tt}C^Ty(t)dt = \int_0^{t_f} e^{A^Tt}C^TCe^{At}dtx_0$$

$$\int_0^{t_f} e^{A^Tt}C^Ty(t)dt = W(t_f)x_0$$

This equation has a unique solution

$$x_0 = W^{-1}(t_f)\int_0^{t_f} e^{A^T t}C^T y(t)dt$$

Proof of Necessity: We want to show that positive definiteness of $W(t_f)$ is a necessary condition for observability over $[0, t_f]$.
Suppose $W(t_f)$ is not positive definite. Then, there is a non-zero vector v such that

$$Ce^{At}v \equiv 0, \, \forall t \in [0, t_f]$$

The output due to $x(0) = x_0$ is $Ce^{At}x_0$ and the out put due to $x(0) = x_0 + v$ is

$$Ce^{At}(x_0 + v) = Ce^{At}x_0 + Ce^{At}v = Ce^{At}x_0$$

The initial states x_0 and $x_0 + v$ produce the same output. Hence, x_0 cannot be uniquely determined from the output. therefore, $W(t_f)$ must be positive definite for the system to be observable. □

Theorem 2.3.2. *The Observability Gramian $W(t_f)$ is positive definite if and only if the $rank\mathcal{O} = n$, where*

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

is the observability matrix.

Proof. $W(t_f)$ is not positive definite if and only if there is a non-zero vector v such that

$$Ce^{At}v \equiv 0, \quad \forall t \in [0, t_f] \quad \Leftrightarrow C\sum_{k=0}^{\infty}\frac{t^k}{k!}A^k v \equiv 0.$$

$$\Leftrightarrow CA^k v = 0, \quad \forall k \geq 0$$

$$\Leftrightarrow CA^k v = 0, \quad \forall k = 0, 1, 2, ..., n-1$$

$$\Leftrightarrow \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} v = 0 \quad \Leftrightarrow \quad rank \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} < n$$

□

Theorem 2.3.3. *Consider an autonomous system $\dot{x} = Ax$ or $\sigma x = Ax$, $y = Cx$*

- *In continuous time let $\tau^* = \epsilon > 0$ be arbitrary.*
- *In discrete time , let $\tau^* = n - 1$, where $n = dim(\mathbf{X})$.*
- $\mathcal{J} = \mathcal{J}(\tau^*) = \ker(W(\tau^*)) = \ker(\mathcal{O})$

Therefore the following are equivalent:

1. *The system is observable*
2. *$W(\tau^*)$ is non singular*
3. *$\mathbf{O}$ has full column rank.*

Moreover in that case, we have reconstruction formulas:

$$x(0) = W(\tau^*)^{-1} \int_0^{\tau^*} e^{A^T t} C^T y(t) dt.$$

or

$$x(0) = W(\tau^*)^{-1} \sum_{i=0}^{\tau^*} (A^T)^i C^T y(i).$$

Proof. We have separately proofed this in the above theorems and hence we don't need to repeat it here. □

Example 2.3.4. *compute $\mathcal{J}(i)$, the set of states that are indistinguishable from zero in time $i = 0, 1, 2, ...$ for the discrete system given by*

$$x(i+1) = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix} x(i), \qquad y(i) = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} x(i)$$

Solution
we have

$$\mathcal{J}(0) = \ker(C) = \ker \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} = Span < \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} > .$$

$$\mathcal{J}(1) = \ker \begin{pmatrix} C \\ CA \end{pmatrix} = \ker \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & -2 \end{bmatrix} = Span < \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} > .$$

$$\mathcal{J}(2) = \ker \begin{pmatrix} C \\ CA \\ CA^2 \end{pmatrix} = \ker \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & -2 \\ 1 & -1 & 4 \end{bmatrix} = \{0\}.$$

Note also that the system is observable.

2.4 Observability Matrix Pairs and Coordinate Transform

In the former section,We have learnt that a matrix pair (A, C) is observable if the associated Kalman observability matrix

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

has full column rank, that is, $rank(O) = n$. If a state space system $\dot{x} = Ax + Bu$, $y = Cx + Du$ is subject to a coordinate transform $x = Tz$, where $T \in \mathbb{R}^{n \times n}$ is invertible, then we get

$$\begin{aligned} \dot{z} &= T^{-1}ATz + T^{-1}Bu \\ y &= CTz + Du \end{aligned}$$

Discrete systems behave analogously. We say that the matrix pair $(T^{-1}AT, CT)$ is similar to the matrix pair (A, C).
In the next theorem we will prove that a coordinate transform does not change structural system properties, mainly observability.

Theorem 2.4.1. *The pair (A, C)is observable if and only if the pair $(T^{-1}AT, CT)$ is observable.*

Proof.

$$\bar{\mathcal{O}} := \mathcal{O}(T^{-1}AT, CT) = \begin{bmatrix} CT \\ CTT^{-1}AT \\ CTT^{-1}A^2T \\ \vdots \\ CTT^{-1}A^{n-1}T \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix} T = \mathcal{O}T$$

The nonsingularity of T implies

$$Rank(\bar{\mathcal{O}}) = Rank(\mathcal{O})$$

which proves that observability is invariant under similarity transform. $\square$

The following result is limited to the single-output case, that is, $p = 1$. Then C is a single row vector. In that case, we simply write c instead of C. The associated Kalman observability matrix is then a square matrix.

Theorem 2.4.2. *Let $A \in \mathbb{R}^{n \times n}$ and $c \in \mathbb{R}^{1 \times n}$, and let (A, c) be an observable matrix pair. Then there exists an invertible matrix $T \in \mathbb{R}^{n \times n}$ such that*

$$\tilde{A} := T^{-1}AT = \begin{bmatrix} 0 & 1 & & \\ \vdots & & \ddots & \\ \vdots & & & \ddots \\ 0 & & & 1 \\ -a_0 & -a_1 & \dots & -a_{n-1} \end{bmatrix} \text{ and } \tilde{c} := cT = [1, 0 ... 0].$$

The numbers a_i are precisely the coefficients of the characteristic polynomial,that is,

$$\chi_A(s) = s^n + a_{n-1}s^{n-1} + ... + a_1 s + a_0.$$

Proof. According to the Hamilton-Cayley theorem, every $n \times n$ matrix A satisfies its own characteristic equation. That means, if

$$\chi_A(\lambda) = det[\lambda I_n - A] = \lambda^n + a_{n-1}\lambda^{n-1} + ... + a_1\lambda + a_0 = 0$$

then,

$$\Rightarrow \chi_A(A) = A^n + a_{n-1}A^{n-1} + ... + a_1 A + a_0 I_n = [0]$$

$$\Rightarrow A^n = -a_{n-1}A^{n-1} - ... - a_1 A - a_0 I_n$$

Thus we have

$$\mathcal{O}A = \begin{bmatrix} cA \\ cA^2 \\ \vdots \\ cA^n \end{bmatrix} = \begin{bmatrix} 0 & 1 & & & \\ \vdots & & \ddots & & \\ \vdots & & & \ddots & \\ 0 & & & & 1 \\ -a_0 & -a_1 & \dots & & -a_{n-1} \end{bmatrix} \begin{bmatrix} c \\ cA \\ \vdots \\ cA^{n-1} \end{bmatrix} = \tilde{A}\mathcal{O}$$

Moreover, we have

$$c = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} c \\ cA \\ \vdots \\ cA^{n-1} \end{bmatrix}$$

Since (A, c) is observable, $\mathcal{O}$ is invertible. Thus we simply put $T = \mathcal{O}^{-1}$. This completes the proof of the theorem. $\square$

If we return to the general multi-output case, we give another result about transforming a given matrix pair into some special form via a similarity transform.

Theorem 2.4.3 (**Kalman Observability Decomposition**). *Let $A \in \mathbb{R}^{n\times n}$ and $C \in \mathbb{R}^{p\times n}$. Let $\mathcal{O}$ be the associated Kalman observability matrix and let $r := rank(\mathcal{O})$. Then there exists an invertible matrix $T \in \mathbb{R}^{n\times n}$ such that*

$$\tilde{A} := T^{-1}AT = \begin{bmatrix} A_1 & 0 \\ A_2 & A_3 \end{bmatrix} \text{ and } \tilde{C} := CT = \begin{bmatrix} C_1 & 0 \end{bmatrix}$$

where $A_1 \in \mathbb{R}^{r\times r}$, $C_1 \in \mathbb{R}^{p\times r}$ and the pair (A_1, C_1) is an observable matrix pair.

Remark 2.3. *The theorem says that via a suitable coordinate transform, namely $x = Tz$, the given system $\dot{x} = Ax$, $y = Cx$ can be put into the form*

$$\begin{aligned} \dot{z}_1 &= A_1 z_1 \\ \dot{z}_2 &= A_2 z_1 + A_3 z_2 \\ y &= C_1 z_1 \end{aligned}$$

where (A_1, C_1) is observable. consider that the output does not depend on z_2 at all. That means there is no way of reconstructing z_2 by measuring the output y. Hence z_2 is not observable. On the other hand, since (A_1, C_1) is observable, $z_1(0)$ is reconstructible. Thus

$$\{z \in \mathbb{R}^n | z \text{ cannot be distinguished from } 0\} = \{0\} \times \mathbb{R}^{n-r}.$$

This can be used to determine the indistinguishable space of the original system,because

$$\{x \in \mathbb{R}^n | x \text{ cannot be distinguished from } 0\} = T(\{0\} \times \mathbb{R}^{n-r})$$

Note that if the original (A, C) is observable, then $r = n$, and the Kalman observability decomposition becomes trivial. Thus the interesting case arises when (A, C) itself is not observable.

Proof. Let k be a $\dim\ker(\mathcal{O})$ corresponding to (A, C) and let $(v_1, v_2, ..., v_r, w_1, w_2, ..., w_{k=n-r})$ is a basis of $\mathbb{R}^n$ where $(w_1, w_2, ..., w_{k=n-r})$ is the basis of $\ker(\mathcal{O})$. Then,

$$T := \begin{bmatrix} V & W \end{bmatrix} = \begin{bmatrix} v_1 & v_2 & \ldots & v_r & w_1 & w_2 & \ldots & w_k \end{bmatrix} \in \mathbb{R}^{n\times n}$$

is invertible matrix. Due to A-invariance of $\ker(\mathcal{O})$, the columns of AW are again in $\ker(\mathcal{O})$. Thus, they can be written as linear combinations of the vectors w_i , that is,

$$AW = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} a_{11}w_1 + a_{12}w_2 + \ldots + a_{1n}w_n \\ a_{21}w_1 + a_{22}w_2 + \ldots + a_{2n}w_n \\ \vdots \\ a_{n1}w_1 + a_{n2}w_2 + \ldots + a_{nn}w_n \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} w_1 + \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} w_2 + \ldots + \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix} w_n$$

$$AW = WA_3$$

for some matrix $A_3 \in \mathbb{R}^{k\times k}$. On the other hand, the columns of AV are in $\mathbb{R}^n$ and thus they can be written as linear combinations of v_i, w_j , that is,

$$AV = VA_1 + WA_2$$

for some matrices $A_1 \in \mathbb{R}^{r\times r}$ and $A_2 \in \mathbb{R}^{k\times r}$. Summing up, we have

$$AT = \begin{bmatrix} AV & AW \end{bmatrix} = \begin{bmatrix} VA_1 + WA_2 & WA_3 \end{bmatrix} = \begin{bmatrix} V & W \end{bmatrix} \begin{bmatrix} A1 & 0 \\ A2 & A3 \end{bmatrix} = T\tilde{A}$$

This gives $T^{-1}AT = \tilde{A}$.
Furthermore, since $\ker(\mathcal{O}) \subseteq \ker(C)$, there exists a matrix $C_1 \in \mathbb{R}^{p\times r}$ such that,

$$CT = \begin{bmatrix} CV & CW \end{bmatrix} = \begin{bmatrix} C_1 & 0 \end{bmatrix}.$$

Now it remains to show the pair (A_1, C_1) is observable. The Kalman observability matrix associated to this matrix pair $(\tilde{A}, \tilde{C})$ is

$$\tilde{O} = \begin{bmatrix} \tilde{C} \\ \tilde{C}\tilde{A} \\ \vdots \\ \tilde{C}\tilde{A}^{(n-1)} \end{bmatrix} = \begin{bmatrix} C_1 & 0 \\ C_1A_1 & 0 \\ \vdots & \\ C_1A_1^{(n-1)} & 0 \end{bmatrix}$$

and this has rank r, just like the Kalman observability matrix of the original matrix pair. Therefore

$$r = rank \begin{bmatrix} C_1 \\ C_1A_1 \\ \vdots \\ C_1A_1^{(n-1)} \end{bmatrix}$$

Due to the Hamilton-Cayley theorem, this implies that

$$r = rank \begin{bmatrix} C_1 \\ C_1 A_1 \\ \vdots \\ C_1 A_1^{(r-1)} \end{bmatrix}$$

which shows that (A_1, C_1) is observable. $\square$

Remark 2.4. *In a Kalman observability decomposition, we clearly have*

$$\chi_A = \chi_{A_1} \cdot \chi_{A_3}$$

and thus

$$spec(A) = spec(A_1) \bigcup spec(A_3).$$

*One calls χ_{A_3} the characteristic polynomial of the unobservable part of A with respect to C, and $\lambda \in spec(A_3)$ an **unobservable mode** of (A, C). There is also a direct way to characterize the unobservable modes of a matrix pair.*

Theorem 2.4.4. *Let $A \in \mathbb{R}^{n\times n}, C \in \mathbb{R}^{p\times n}$, and $\lambda \in \mathbb{C}$. Then the following are equivalent:*

1. *λ is an unobservable mode of (A, C)*
2. *$rank \begin{bmatrix} \lambda I - A \\ C \end{bmatrix} < n$.*

Proof. From Kalman observability decomposition we have

$$A = \begin{bmatrix} A_1 & 0 \\ A_2 & A_3 \end{bmatrix} \qquad \text{and} \qquad C = \begin{bmatrix} C_1 & 0 \end{bmatrix}$$

where (A_1, C_1) is observable. Then we have

$$H(\lambda) := \begin{bmatrix} \lambda I - A \\ C \end{bmatrix} = \begin{bmatrix} \lambda I - A_1 & 0 \\ -A_2 & \lambda I - A_3 \\ C_1 & 0 \end{bmatrix}$$

If λ is an unobservable mode, then it is an eigenvalue of A_3. Thus it makes $\lambda I - A_3$ singular, Showing that the column rank of $H(\lambda)$ cannot be full, it must be less than n. Conversely, assume that the rank of $H(\lambda)$ is not full. Then there exists a vector $x \neq 0$ such that $H(\lambda)x = 0$, that is,

$$\begin{gathered} (\lambda I - A_1)x_1 = 0 \\ -A_2 x_1 + (\lambda I - A_3)x_2 = 0 \\ C_1 x_1 = 0 \end{gathered}$$

From the first and the third equation we see that

$$C_1 x_1 = 0, \quad C_1 A_1 x_1 = 0, \quad C_1 A_1^2 x_1 = 0, ...$$

The observability of (A_1, C_1) yeilds that $x_1 = 0$. Then $x_2 \neq 0$ and

$$(\lambda I - A_3)x_2 = 0.$$

This implies that λ is an eigenvalue of A_3. That means, it is an unobservable mode of (A, C). $\square$

As a direct consequence of this, we obtain another characterization of observable matrix pairs.

Corollary 2.3 (**Popov - Belevich- Hautus Test**). *The following are equivalent:*

1. *(A, C) is observable.*

2. *There is no non-zero vector x such that $Ax = \lambda x$, $Cx = 0$*

Proof. $(1 \Rightarrow 2)$
Suppose $\exists x \neq 0$ such that $Ax = \lambda x$, $Cx = 0$

$$\begin{aligned}
&\Rightarrow CAx = \lambda Cx = 0 \\
&\Rightarrow CA^2x = \lambda CAx = 0 \\
&\quad . \\
&\quad . \\
&\quad . \\
&\Rightarrow CA^{n-1}x = \lambda CA^{n-1}x = 0 \\
&\Rightarrow \mathcal{O}x = 0
\end{aligned}$$

$\Rightarrow$ $Rank(\mathcal{O}) < n \Rightarrow$ the pair (A,C) is not observable. Hence, it must be the case that $\nexists x \neq 0$ such that $Ax = \lambda x$, $Cx = 0$.

$(2 \Rightarrow 1)$

Assume that there is no non-zero vector x such that $Ax = \lambda x$, $Cx = 0$. By the same derivation we used as in the above we get

$$\nexists x \neq 0 \text{ Such that } \mathcal{O}(A, C)x = 0$$

$$\Rightarrow Rank(\mathcal{O}(A, C)) = n$$

Where n is the order of A. Hence The pair (A, C) is observable. $\square$

Corollary 2.4 (**Hautus test for observability**). *The following are equivalent:*

1. *(A, C) is observable.*

2. *The matrix $H(\lambda) = \left[\begin{smallmatrix} \lambda I - A \\ C \end{smallmatrix}\right]$ has full column rank for all $\lambda \in \mathbb{C}$.*

The polynomial matrix $H = \begin{bmatrix} \lambda I - A \\ C \end{bmatrix} \in \mathbb{R}[s]^{(n+p)\times n}$ is called **Hautus observability matrix**. The proof of this theorem follows directly from the above corollary 2.3 and we don't have to prove it here.

2.5 Asymptotic observability

Observability means

$$x \in \mathcal{J} \Rightarrow x = 0.$$

An alternative formulation is

$$\eta(t, 0, x, 0) = 0 \ \forall t \geq 0 \Rightarrow \varphi(t, 0, x, 0) = 0 \ \forall t \geq 0.$$

For asymptotic observability, one is satisfied if this is true in the limit as $t \to \infty$.

Definition 2.5.1. *We say that a state space system $\dot{x} = Ax$ or $\sigma x = Ax$, $y = Cx$, is **asymptotically observable** if $\eta(t, 0, x, 0) = 0 \ \forall t \geq 0$ implies that*

$$\lim_{t\to\infty} \varphi(t, 0, x, 0) = 0.$$

Clearly, observability implies asymptotic observability.

Theorem 2.5.1. *A state space system is asymptotically observable if and only if its unobservable modes λ are asymptotically stable, that is, $Re(\lambda) < 0$ in continuous time, and $|\lambda| < 1$ in discrete time.*

2.6 Observable latent variable descriptions

The Hautus test gives us an idea about how to generalize the notion of observability from state space systems $\dot{x} = \mathbf{A}x + \mathbf{B}u, y = \mathbf{C}x + \mathbf{D}u$ to general systems $\mathbf{R}(\frac{d}{dt})w = \mathbf{M}(\frac{d}{dt})l$ where $\mathbf{R} \in \mathbb{R}[s]^{p\times q}$, $\mathbf{M} \in \mathbb{R}[s]^{p\times r}$ and $w \in \mathbf{A}^q$, $l \in \mathbf{A}^r$. In a state space system,

$$\mathbf{R} = \begin{bmatrix} \mathbf{B} & 0 \\ -\mathbf{D} & I \end{bmatrix}, \ \mathbf{M} = \begin{bmatrix} sI - \mathbf{A} \\ \mathbf{C} \end{bmatrix} \text{ and } w = \begin{bmatrix} u \\ y \end{bmatrix}, \ l = x.$$

The polynomial matrix $\mathbf{M}$ is recognized as the Hautus observability matrix.
In this section, we restrict our discussion to continuous systems, that is, $\mathcal{A} = \mathcal{D}'(\mathbf{T})$, where $\mathbf{T} = \mathbb{R}, \mathbb{R}_+$.

Definition 2.6.1. *We say that the latent variables l can be observed from the manifest variables w if l is uniquely determined by w, which means that*

$$w_1 = w_2 \Rightarrow l_1 = l_2.$$

Theorem 2.6.1. *(**Generalized Hautus test**) Let $\mathcal{A} = \mathcal{D}'(T)$ for $\mathbf{T} = \mathbb{R}$ or $\mathbb{R}_+$ and let $\mathbf{M} \in \mathbb{R}[s]^{p\times r}$. Without loss of generality, let $\mathbf{M}$ have full column rank. Then the latent variable description of*

$$\mathcal{B} = \{w \in \mathcal{A}^q | \exists l \in \mathcal{A}^r : \mathbf{R}(\frac{d}{dt})w = \mathbf{M}(\frac{d}{dt})l\}$$

is observable if and only if rank$(\mathbf{M}(\lambda)) = r$ for all $\lambda \in \mathbb{C}$.

Proof. Observe

$$\begin{aligned} \dot{x} = \mathbf{A}x + \mathbf{B}u \qquad &\Leftrightarrow \qquad (sI - \mathbf{A})x = \mathbf{B}u \\ y = \mathbf{C}x + \mathbf{D}u \qquad & \qquad\qquad \mathbf{C}x = -\mathbf{D}u + y \end{aligned}$$

$$\Leftrightarrow \qquad \begin{bmatrix} sI-\mathbf{A} \\ \mathbf{C} \end{bmatrix} x = \begin{bmatrix} B & 0 \\ -D & I \end{bmatrix} \begin{bmatrix} u \\ y \end{bmatrix}$$

$$\Leftrightarrow \qquad Mx = Rw \text{ where } w = \begin{bmatrix} u \\ y \end{bmatrix}, M = \begin{bmatrix} sI-\mathbf{A} \\ \mathbf{C} \end{bmatrix} \text{ and } R = \begin{bmatrix} B & 0 \\ -D & I \end{bmatrix}$$

From Hautus test, the state space system is observable

$$\Leftrightarrow \qquad \text{The matrix } \begin{bmatrix} \lambda I-\mathbf{A} \\ \mathbf{C} \end{bmatrix} \text{ has full column rank for all } \lambda \in \mathbb{C}$$

$$\Leftrightarrow \qquad \text{The matrix } M(\lambda) \text{ has full column rank for all } \lambda \in \mathbb{C}$$

$$\Leftrightarrow Rank(M(\lambda)) = r \text{ for all } \lambda \in \mathbb{C}$$

Thus, the theorem! □

Appendices

Derivation of state-space representation

Theorem :
LetR $\in \mathbb{R}[s]^{p\times q}$. There exists an integer $n \in \mathbb{N}$ and real matrices $K \in \mathbb{R}^{n\times n}$, $L \in \mathbb{R}^{n\times q}$, $M \in \mathbb{R}^{p\times n}$, $N \in \mathbb{R}^{p\times q}$ such that the system law $R(\frac{d}{dt})w = 0$ has a first order latent variable representation of the form

$$R(\frac{d}{dt})w = 0 \quad \Leftrightarrow \quad \exists x \in \mathcal{A}^n : \begin{cases} \frac{d}{dt}x = Kx + Lw \\ 0 = Mx + Nw \end{cases}$$

and similarly for σ instead of $\frac{d}{dt}$.

Proof. Let $R = R_d s^d + ... + R_1 s + R_0$ with $R_i \in \mathbb{R}^{p\times q}$. We put $n = dp$ and

$$K = \begin{bmatrix} 0 & \cdots & & 0 \\ I_p & \ddots & & \vdots \\ & \ddots & \ddots & \vdots \\ & & I_p & 0 \end{bmatrix} \in \mathbb{R}^{n\times n}, \;\; L = \begin{bmatrix} R_0 \\ \vdots \\ R_{d-1} \end{bmatrix} \in \mathbb{R}^{n\times p}, \;\; M = \begin{bmatrix} 0 & \cdots & I_p \end{bmatrix} \in \mathbb{R}^{p\times n}, \;\; N = R_d \in \mathbb{R}^{p\times q}.$$

Then,

$$\begin{bmatrix} \frac{d}{dt}I_n - K \\ -M \end{bmatrix} x = \begin{bmatrix} L \\ N \end{bmatrix} w$$

can be pre-multiplied by $U(\frac{d}{dt})$, where U is the unimodular matrix of size $n+p = (d+1)p$

$$U = \begin{bmatrix} I_p & sI_p & \cdots & s_d I_p \\ & \ddots & \ddots & \vdots \\ & & \ddots & \vdots \\ & & & sI_p \\ & & & I_p \end{bmatrix}$$

to obtain the equivalent equation

$$\begin{bmatrix} 0 \\ -I_n \end{bmatrix} x = \begin{bmatrix} R \\ * \end{bmatrix} (\frac{d}{dt})w$$

where the * denotes a polynomial matrix whose precise form is not important here, because we only need that such an x exists if and only if $R(\frac{d}{dt})w = 0$. □

According to this theorem, for every input-output representation $P(\frac{d}{dt})y = Q(\frac{d}{dt})u$, we can find, , a first order representation of the form

$$P(\frac{d}{dt})y = Q(\frac{d}{dt})u \quad \Leftrightarrow \quad \exists x \in \mathcal{A}^n : \begin{cases} \frac{d}{dt}x = Kx + L_1 u + L_2 y \\ 0 = Mx + N_1 u + N_2 y \end{cases}$$

and similarly for σ instead of $\frac{d}{dt}$. Then we can solve the second equation for y to obtain

$$y = -N_2^{-1}(Mx + N_1 u)$$

and plug that into the first equation. We get

$$\frac{d}{dt}x = (K - L_2N_2^{-1}M)x + (L_1 - L_2N_2^{-1}N_1)u.$$

Setting

$$\begin{aligned} A &= K - L_2N_2^{-1}M \\ B &= L_1 - L_2N_2^{-1}N_1 \\ C &= -N_2^{-1}M \\ D &= -N_2^{-1}N_1 \end{aligned}$$

we have

$$P(\frac{d}{dt})y = Q(\frac{d}{dt})u \quad \Leftrightarrow \quad \exists x \in \mathcal{A}^n : \begin{cases} \frac{d}{dt}x = Ax + Bu \\ y = Cx + Du \end{cases}$$

The explicit equations

$$\begin{aligned} \dot{x} &= Ax + Bu \qquad or \qquad & \sigma x &= Ax + Bu \\ y &= Cx + Du & y &= Cx + Du \end{aligned}$$

are called **state space representations**.

A- invariant subspace

Definition B_1**:**(A-***invariant subspace***)
Let $A \in \mathbb{R}^{n\times n}$ and let $\mathcal{V}$ be a linear subspace of $\mathbb{R}^n$. We call $\mathcal{V}$ an A-invariant subspace if for all $v \in \mathcal{V}$, $Av \in \mathcal{V}$. Notation: $A\mathcal{V} \subseteq \mathcal{V}$.

Theorem B_1**:**
The kernel of $\mathcal{O}$ is the largest A-invariant subspace contained in the kernel of C. Therefore, a system defined by state space equation is observable if and only if $\mathbf{0}$ is the largest A-invariant subspace contained in $\ker(C)$.

Proof. Choose $x \in \ker(\mathcal{O})$. Then

$$Cx = CAx = ... = CA^{n-1}x = 0.$$

By the CayleyHamilton theorem, it follows that also $CA^nx = 0$, and hence $Ax \in \ker(\mathcal{O})$, which implies that $\ker(\mathcal{O})$ is A-invariant.

Furthermore, $\mathcal{O}x = 0$ implies that $Cx = 0$, and hence $\ker(\mathcal{O}) \subseteq \ker(C)$. Therefore,

$\ker(\mathcal{O})$ is an A-invariant subspace contained in $\ker(C)$.

To show that it is the largest such subspace, assume that $\mathcal{V}$ is an A-invariant subspace contained in the kernel of C. Choose $x \in \mathcal{V}$. Then, since $\mathcal{V}$ is A-invariant, also

$$Ax\ ,A^2x\ ,\ldots,\ A^{n-1}x \in \mathcal{V},$$

and since $\mathcal{V}$ is contained in $\ker(C)$, we conclude that

$$Cx = CAx = \ = CA^{n-1}x = 0.$$

This implies that $x \in \ker(O)$, and hence $\mathcal{V} \subseteq \ker(\mathcal{O})$, as claimed. □

Bibliography

[1] Eva Zerz; A lecture note on *Introduction to Systems and Control Theory*,Lehrstuhl D für Mathematik, RWTH Aachen, 52062 Aachen, Germany .

[2] Kalman, R. E; On the *General Theory of Control Systems*, IRE Transactions on Automatic Control, Volume 4, Issue 3, p110, 1959. ISSN 0096199X

[3] Bijoy K, Ghosh and Joachim Rosenthal(IEEE Transactions on Automatic control, Vol. 40, No. I , 1995)

[4] Prof. Khan; A lacture note on *Advanced control: Controllability and Observability*, Spring 2013 - EE 194,

[5] J. W. Polderman, J. C. Willems; *Introduction to Mathematical Systems Theory*, Texts in Applied Mathematics 26, Springer, 1998.