university of groningen
faculty of science and engineering
mathematics and applied mathematics

Positive Linear Systems

Bachelor's Project Applied Mathematics
July 2024
Student: N.T. Nguyen
First supervisor: dr. ir. H.J. van Waarde
Second assessor: prof. M.K. Camlibel

Abstract

This thesis serves as an introduction to "Positive Systems Theory" by providing an overview of the fundamental results concerning positive systems. Generally speaking, a dynamical system is positive if its describing variables only take nonnegative values. Motivated by various real-life examples, we study linear time-invariant systems in continuous and in discrete time and we provide characterisations for their positivity. Next, we show several ways to verify whether a positive system is asymptotically stable, which reveal that checking stability properties of large-scale systems remains practically feasible. We proceed with the widely recognised Kalman-Yakubovich-Popov lemma adapted to positive systems, which bridges the frequency domain and state space. Then, we tackle the positive stabilisation problem, i.e. we seek to design a controller, which achieves both stability and positivity. Finally, we take first steps towards data-driven analysis and control of positive systems by solving the problem of system identification. Throughout this thesis, we demonstrate our results with numerical examples.

Table of Contents

Acknowledgements

First and foremost, I would like to thank my main supervisor, dr. ir. Henk J. van Waarde. Professor van Waarde, it has been an immense honour and pleasure working alongside you for the past two years. Doing research with you has been one of the most "positive" and "informative" experiences during the Bachelor's programme. Thank you for your limitless patience, the great and abundant feedback, the opportunity to attend ECC24, the enthusiasm, the stories and the jokes. Thank you also for never lowering your standards and making me want to become a better writer and storyteller.

I would like to extend my gratitude to prof. dr. ir. Bart Besselink and prof. M. Kanat Camlibel. Thank you for introducing me to and deepening my interest in systems and control theory.

Nothing from the past three years would have been possible without the endless support of my father, mother and brother.

Finally, I would like to thank Teya, Hristo, Yoko, Jijo, Mirinski, Tuni and Martin for being great friends. May we reach gold rank in Siege soon.

Chapter 1

Introduction

In this thesis, we are interested in a particular class of dynamical systems, namely, those, whose state variables, inputs and outputs only take nonnegative values. There are ample examples based on applications of these so-called positive systems, some of which include models in social sciences, epidemiology, biology, pharmacology, biochemical engineering, econometrics and stochastic processes. In fact, population density, concentrations, prices and probabilities are variables with a natural non-negativity constraint.

There is a vast literature on "Positive Systems Theory" that spans over the past 45 years. If we include the contributions from "Positive Matrix Theory", for instance, the renowned Perron-Frobenius theorem, then we can even say over the past 120 years. Regardless, we attribute the beginning of positive systems theory to David G. Luenberger. In [1, Chapter 6], he studies the positivity, equilibria and stability of continuous- and discrete-time linear time-invariant (LTI) systems with constant inputs. Additionally, the interested reader should refer to the books by Abraham Berman and Robert J. Plemmons [2], Lorenzo Farina and Sergio Rinaldi [3] as well as Tadeusz Kaczorek [4]. The tutorial paper [5] is a good starting point for those who seek a shorter introduction.

The goal of this thesis is to provide an overview of the foundational results for positive LTI systems in continuous and discrete time. We, then, prove them in an accessible manner. In the last part of the thesis, we expand the literature on positive systems by moving to a data-based setting. Furthermore, there are several numerical examples, which illustrate the results. In what follows, we discuss the structure of the thesis in more detail.

In Chapter 2, we introduce a model in continuous and in discrete time. We express interest in the conditions under which they are positive and asymptotically stability. The next chapter is dedicated to investigating these system properties.

Chapter 3 opens with a formal definition of a positive system, after which we check when an LTI system is positive. Then, we provide tests for asymptotic stability, which show that positive systems remain tractable as the system dimensions increase. Moreover, they reveal a method to reformulate certain problems, so that they have a lower computational cost. We also discuss their application to Lyapunov stability theory. To exemplify the main statements, we verify whether numerical examples of the aforementioned models are positive and asymptotically stable.

Under the assumption of positivity and asymptotic stability, we continue with the celebrated Kalman-Yakubovich-Popov (KYP) lemma in Chapter 4, which we adapt to positive systems. This is a cornerstone result in modern control theory, which connects the frequency domain and state space. Meaningful differences arise between the version we state and the one for unconstrained systems. We prove the KYP lemma with the help of duality theory for linear matrix inequalities.

When a system is neither positive nor asymptotically stable, we seek to design a state feedback controller, which achieves both. We devote Chapter 5 to solving the positive stabilisation problem.

We leverage the results in Chapter 3 in order to derive necessary and sufficient conditions for the existence of such a controller by means of semidefinite programming. Subsequently, we restate the conditions in terms of linear programming.

In contrast to the previous chapters, in Chapter 6, we shift our attention to data-driven analysis and control. This has been a recently emerging paradigm, inspired by the rise in complexity of modern systems as well as the wide availability of data. Research has been dedicated to developing data-based methods, which allow us to work with systems without going through the modelling process. One of the approaches to data-driven control involves the intermediate step of system identification. In Chapter 6, we consider an unknown positive system and assume that we have access to data, which are generated by the system. We state necessary and sufficient conditions under which the positive system can be uniquely identified from the data.

In Chapter 7, we recapitulate the main results in the thesis. We conclude with a list of possible future research topics.

1.1 Notation

Let $\mathbf{e}_i$ be the ith *standard basis vector* in $\mathbb{R}^n$ and $\mathbb{1}_n$ be the vector of length n, whose entries are all equal to 1. Consider a matrix $A \in \mathbb{R}^{n\times n}$ and a vector $\mathbf{a} \in \mathbb{R}^n$. We indicate the ith element of $\mathbf{a}$ by a_i. The set $\sigma(A)$, called the *spectrum* of A, is the one containing the eigenvalues of A and the *spectral radius* of A is defined as $\rho(A) := \max\{|\lambda| \mid \lambda \in \sigma(A)\}$. We denote the *trace* of A by $\mathrm{tr}(A)$. We define $\mathrm{diag}(A)$ as the vector containing the diagonal elements of the matrix A. Conversely, we define $\mathrm{diag}(\mathbf{a})$ as the diagonal matrix, whose elements correspond to the vector $\mathbf{a}$.

We denote by $\mathbb{S}^n$ the *set of real symmetric $n \times n$ matrices*. Let $A \in \mathbb{S}^n$. If $\mathbf{x}^T A\mathbf{x} > 0$ for all nonzero $\mathbf{x} \in \mathbb{R}^n$, then we call A *positive definite*, denoted by $A \succ 0$. If instead $\mathbf{x}^T A\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$, then we call it *positive semidefinite*, denoted by $A \succeq 0$. *Negative definiteness* ($A \prec 0$) and *negative semidefiniteness* ($A \preceq 0$) are defined similarly. By $A \succ B$, where $B \in \mathbb{S}^n$, we mean that $A - B \succ 0$. The meaning of $A \succeq B$, $A \prec B$ and $A \preceq B$ is similar.

We use $\mathbb{Z}_+$ and $\mathbb{R}_+$ to denote the *set of nonnegative integers* and the *set nonnegative real numbers*, respectively. We call a matrix $A \in \mathbb{R}^{n\times m}$ and a vector $\mathbf{a} \in \mathbb{R}^n$ *nonnegative*, if all of their elements are nonnegative. We denote this by $A \geq 0$ and $\mathbf{a} \geq 0$. If all of their elements are positive instead, then we call them *positive* and denote this by $A > 0$ and $\mathbf{a} > 0$. If $-A \geq 0$, $-\mathbf{a} \geq 0$, $-A > 0$ and $-\mathbf{a} > 0$, then we denote this by $A \leq 0$, $\mathbf{a} \leq 0$, $A < 0$ and $\mathbf{a} < 0$. For all $A, B \in \mathbb{R}^{n\times m}$ and $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, by $A \geq B$ and $\mathbf{a} \geq \mathbf{b}$ we mean that $A - B \geq 0$ and $\mathbf{a} - \mathbf{b} \geq 0$. The meaning of $A - B > 0$, $\mathbf{a} - \mathbf{b} > 0$, $A - B \leq 0$, $\mathbf{a} - \mathbf{b} \leq 0$, $A - B < 0$ and $\mathbf{a} - \mathbf{b} < 0$ is similar.

For a matrix $A \in \mathbb{C}^{n\times m}$, A^T is its *transpose* and A^* is its *conjugate transpose*. We indicate its ith row by $A_{i\bullet}$, its ith column by $A_{\bullet i}$ and its (i, j)th element by a_{ij}. We denote the real part of a complex number $\lambda \in \mathbb{C}$ by $\mathrm{Re}(\lambda)$.

Chapter 2

Motivation

We begin by motivating the study of positive systems in both continuous (Section 2.1) and discrete time (Section 2.2) with the help of two specific models, namely, radioactive decay and the Leslie model. For each one, we ask two very natural questions, which determine the first two topics positive systems theory starts off with.

2.1 Radioactive Decay

The nuclei of an unstable isotope decay over time. For a radioactive isotope, it may happen that the result of the decay is not itself stable and this can be modelled as a chain of reactions [6, Example 2.4.5], which ends with a stable isotope. More specifically, consider the Uranium-238 decay chain

$$
\begin{gathered}
{}^{238}\mathrm{U} \xrightarrow{\lambda_1} {}^{234}\mathrm{Th} \xrightarrow{\lambda_2} {}^{234}\mathrm{Pa} \xrightarrow{\lambda_3} {}^{234}\mathrm{U} \xrightarrow{\lambda_4} {}^{230}\mathrm{Th} \xrightarrow{\lambda_5} {}^{226}\mathrm{Ra} \xrightarrow{\lambda_6} {}^{222}\mathrm{Rn} \xrightarrow{\lambda_7} {}^{218}\mathrm{Po} \cdots \\
\cdots \xrightarrow{\lambda_8} {}^{214}\mathrm{Pb} \xrightarrow{\lambda_9} {}^{214}\mathrm{Bi} \xrightarrow{\lambda_{10}} {}^{214}\mathrm{Po} \xrightarrow{\lambda_{11}} {}^{210}\mathrm{Pb} \xrightarrow{\lambda_{12}} {}^{210}\mathrm{Bi} \xrightarrow{\lambda_{13}} {}^{210}\mathrm{Po} \xrightarrow{\lambda_{14}} {}^{206}\mathrm{Pb},
\end{gathered}
$$

where $\lambda_i > 0$ $(i = 1, \ldots, 14)$ is the decay rate of the nuclei of the ith isotope. If we assume that the decay rate is proportional to the nuclei density, then we can mathematically describe the number of nuclei for each isotope in the chain as the system of ordinary differential equations

$$
\begin{aligned}
\dot{x}_1(t) &= -\lambda_1 x_1(t), \\
\dot{x}_2(t) &= \lambda_1 x_1(t) - \lambda_2 x_2(t), \\
\dot{x}_2(t) &= \lambda_2 x_2(t) - \lambda_3 x_3(t), \\
&\vdots \\
\dot{x}_{14}(t) &= \lambda_{13} x_{13}(t) - \lambda_{14} x_{14}(t), \\
\dot{x}_{15}(t) &= \lambda_{14} x_{14}(t),
\end{aligned}
$$

which can be rewritten as a continuous-time linear system:

$$
\dot{\mathbf{x}}(t) = \Lambda \mathbf{x}(t), \tag{2.1}
$$

$$
\Lambda := \begin{bmatrix} -\lambda_1 & & & & \\ \lambda_1 & -\lambda_2 & & & \\ & \lambda_2 & -\lambda_3 & & \\ & & \ddots & \ddots & \\ & & & \lambda_{13} & -\lambda_{14} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{14} \end{bmatrix}.
$$

We remark that (2.1) excludes the dynamics of Lead-206, which is a stable isotope, as we are interested in the behaviour of the unstable nuclei. Because radioactive decay is a natural process, firstly, we express interest in the following question:

1. Does the number of nuclei stay nonnegative, i.e. under what conditions does $\mathbf{x}$ remain nonnegative for all times?

Secondly, in the context of decay, we can ask

2. Do the unstable nuclei decay fully, i.e. under what conditions does $\mathbf{x}$ converge to 0 as $t \to \infty$?

2.2 Leslie Model

A common example of positive systems are population models – population density is a variable with a nonnegativity constraint. In particular, we explain the Leslie model [3, Chapter 13], which is concerned with age-structured populations. For such populations, fertility and survival rates strongly depend on age. The Leslie model calculates the population level yearly (or per reproduction season), which implies that it is a discrete-time model. Suppose that the survival rates of males and females are the same and that the sex ratio is balanced. Then, we can express the number of females at age $i = 1, \dots, n$ as

$$x_i(t+1) = s_{i-1} x_{i-1}(t),$$

where $s_j \geq 0$ ($j = 0, \dots, n-1$) denotes the survival rate at age j. Let $f_k \geq 0$ (k = 1, ..., n) be the fertility rate of females at age k. The number of newborns at time t is

$$x_0(t) = f_1 x_1(t) + \dots + f_n x_n(t).$$

We can describe the population dynamics as a discrete-time linear system:

$$\mathbf{x}(t+1) = L\mathbf{x}(t),$$

where L, called the Leslie matrix, is

$$L := \begin{bmatrix} s_0 f_1 & s_0 f_2 & \cdots & s_0 f_{n-1} & s_0 f_n \\ s_1 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & s_{n-1} & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}.$$

We can extend the model to include exogenous inputs $b_i u(t) = s_{i-1} u(t)$ from immigration or stocking of a population of age i as well as an output $\mathbf{y}$ that serves as a population indicator, where the type of indicator depends on the choice of the vector $\mathbf{c}$:

$$\begin{aligned} \mathbf{x}(t+1) &= L\mathbf{x}(t) + \mathbf{b}u(t), \\ \mathbf{y}(t) &= \mathbf{c}^T \mathbf{x}(t). \end{aligned} \tag{2.2}$$

For system (2.2) we are interested in the question

1. Are the population levels and indicator always nonnegative, i.e. under what conditions do $\mathbf{x}$ and $\mathbf{y}$ remain nonnegative for all times?

Note that, here, the survival and fertility rates are constant. Many factors, e.g. lack of food and epidemics (and modern medicine, in the case of human beings), make it hard to calculate survival and fertility rates accurately when looking at large time horizons. For this reason, the Leslie model is best suited for short-term forecasts and low population species, possibly doomed to extinction:

2. Does the species become extinct at some point, i.e. under what conditions does $\mathbf{x}$ converge to 0 (as $t \to \infty$)?

Chapter 3

Positive Systems

In this chapter, we answer the questions posed in Chapter 2, starting with a formal definition of the concept of positivity. Section 3.1 is dedicated to answering the questions concerning the positivity of systems. Section 3.2 studies the long-term stability of positive systems. The results in that section lead to important implications, which are discussed in Subsection 3.2.2. Lastly, we use numerical examples of both models in Chapter 2 to illustrate the main results.

We are interested in the following two linear time-invariant systems:

$$\Sigma_{\mathrm{c}}\colon \begin{cases} \dot{\mathbf{x}}(t) = A\mathbf{x}(t) + B\mathbf{u}(t) \\ \mathbf{y}(t) = C\mathbf{x}(t) + D\mathbf{u}(t) \end{cases}, \ t \in \mathbb{R}_+ \qquad \Sigma_{\mathrm{d}}\colon \begin{cases} \mathbf{x}(t+1) = A\mathbf{x}(t) + B\mathbf{u}(t) \\ \mathbf{y}(t) = C\mathbf{x}(t) + D\mathbf{u}(t) \end{cases}, t \in \mathbb{Z}_+$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$, $D \in \mathbb{R}^{p\times m}$ and $(\mathbf{u}, \mathbf{x}, \mathbf{y}) \in \mathbb{R}^{m+n+p}$. Informally, the variables of a positive system are subject to a positivity (or nonnegativity) constraint. The literature on positive systems theory, however, differentiates between two notions of positivity.

Definition 3.1. Systems Σ_{c} and Σ_{d} are called *internally positive* if, for every initial condition $\mathbf{x}(0) \geq 0$ and every input $\mathbf{u} \geq 0$, the state $\mathbf{x}$ and output $\mathbf{y}$ remain nonnegative for all times.

Definition 3.2. Systems Σ_{c} and Σ_{d} are called *externally positive* if the state $\mathbf{x}$ and output $\mathbf{y}$ remain nonnegative for all times when the initial condition is $\mathbf{x}(0) = 0$ and every input $\mathbf{u}$ is nonnegative.

Evidently, internal positivity implies external positivity. In the rest of the thesis, we study the former, so we refer to internally positive systems as positive systems.

3.1 Characterisation

In continuous-time models with a positivity constraint on the variables, typically, a so-called Metzler[1] matrix with a special structure appears.

Definition 3.3. A real square matrix is called *Metzler* if its off-diagonal elements are nonnegative.

Now, we present and prove the following result, which characterises positive systems.

Theorem 3.4. *Consider systems Σ_c and Σ_d. Then,*

1. *Σ_c is positive if and only if A is Metzler and B, C, D are nonnegative;*
2. *Σ_d is positive if and only if A, B, C and D are nonnegative.*

[1] Named after the American economist Lloyd A. Metzler (1913-1980). His main contributions lie in international trade theory, money, interest and prices, business cycles and economic fluctuations as well as mathematical economics and statistics.

Proof. **(1|⇒):** Assume Σ_{c} is positive. Firstly, let the system be subject to zero inputs at all times. If we take $\mathbf{e}_i$ as an initial condition, then

$$\dot{\mathbf{x}}(0) = A\mathbf{e}_i = A_{\bullet i},$$

i.e. the rate of change of the elements of $\mathbf{e}_i$ is determined by the ith column of A. Because the system is positive, the derivatives at the zero elements of $\mathbf{e}_i$ need to be nonnegative. In other words, we need that $A_{ji} \geq 0$ for all $j \neq i$. Repeating this argument for all $i = 1, \ldots, n$, by Definition 3.3, A is Metzler. Looking at the output equation, in order for

$$\mathbf{y}(0) = C\mathbf{x}(0)$$

to be nonnegative, given an arbitrary $\mathbf{x}(0) \geq 0$, we need that $C \geq 0$.

Secondly, take the zero initial condition. For an arbitrary input $\mathbf{u}(0) \geq 0$, we have

$$\begin{aligned}\dot{\mathbf{x}}(0) &= B\mathbf{u}(0),\\ \mathbf{y}(0) &= D\mathbf{u}(0).\end{aligned}$$

By the same rate of change and output nonnegativity arguments, we conclude that $B \geq 0$ and $D \geq 0$.

(1|⇐): Conversely, suppose A is Metzler and B, C, D are nonnegative. It is clear from the general solution of the system

$$\begin{aligned}\mathbf{x}(t) &= e^{At}\mathbf{x}(0) + \int_0^t e^{A(t-s)}B\mathbf{u}(s)\,ds,\\ \mathbf{y}(t) &= Ce^{At}\mathbf{x}(0) + C\int_0^t e^{A(t-s)}B\mathbf{u}(s)\,ds + D\mathbf{u}(t)\end{aligned}$$

that we only need to make sure that the matrix exponential of a Metzler matrix is nonnegative. Because of the structure of A, we can choose a sufficiently large $\alpha > 0$ such that $A + \alpha I \geq 0$. By using the definition of matrix exponential, note that

$$\begin{aligned}e^{(A+\alpha I)t} &:= \sum_{k=0}^{\infty} \frac{(A+\alpha I)^k t^k}{k!} \geq 0,\\ e^{(-\alpha I)t} &= \begin{bmatrix} e^{-\alpha t} & & \\ & \ddots & \\ & & e^{-\alpha t}\end{bmatrix} \geq 0.\end{aligned}$$

Lastly, by [7, Corollary 15.1.6], it follows that

$$e^{(A+\alpha I)t}e^{(-\alpha I)t} = e^{At} \geq 0.$$

(2|⇒): Let Σ_{d} be positive. We can again assess the properties of A and C by selecting zero inputs as well as those of B and D if the initial condition is $\mathbf{x}(0) = 0$:

$$\left.\begin{array}{l}\mathbf{x}(1) = A\mathbf{x}(0),\\ \mathbf{y}(0) = C\mathbf{x}(0),\\ \text{arbitrary } \mathbf{x}(0) \geq 0\end{array}\right\} \Rightarrow A \geq 0, C \geq 0, \qquad \left.\begin{array}{l}\mathbf{x}(1) = B\mathbf{u}(0),\\ \mathbf{y}(0) = D\mathbf{u}(0),\\ \text{arbitrary } \mathbf{u}(0) \geq 0\end{array}\right\} \Rightarrow B \geq 0, D \geq 0.$$

(2|⇐): The reverse implication is immediate from the general solution of the system:

$$\begin{aligned}\mathbf{x}(t) &= A^t\mathbf{x}(0) + \sum_{k=0}^{t-1} A^{t-(k+1)}B\mathbf{u}(k),\\ \mathbf{y}(t) &= CA^t\mathbf{x}(0) + C\sum_{k=0}^{t-1} A^{t-(k+1)}B\mathbf{u}(k) + D\mathbf{u}(t).\end{aligned}$$

■

Consider again the decay chain system (2.1) and the Leslie model (2.2). Clearly, Λ is Metzler and $L \geq 0$. The physical meaning of the term $\mathbf{b}u(t)$ implies that it always has a nonnegative contribution to the dynamics, i.e. $\mathbf{b} \geq 0$. A common population indicator is the total population, which can be modelled by selecting $\mathbf{c} = \mathbb{1}_n > 0$. Therefore, as expected, both systems are positive.

3.2 Asymptotic Stability

We turn our attention to the stability property of positive systems. Consider the autonomous systems

$$\Sigma_{\mathrm{c}}^{\mathrm{auto}} \colon \dot{\mathbf{x}}(t) = A\mathbf{x}(t), \qquad\qquad \Sigma_{\mathrm{d}}^{\mathrm{auto}} \colon \mathbf{x}(t+1) = A\mathbf{x}(t),$$

which are the systems Σ_{c} and Σ_{d} if the zero input is applied. We define asymptotic stability as follows.

Definition 3.5. The positive systems $\Sigma_{\mathrm{c}}^{\mathrm{auto}}$ and $\Sigma_{\mathrm{d}}^{\mathrm{auto}}$ are called *asymptotically stable* if, for every nonnegative initial condition, the state evolution $\mathbf{x}$ converges to 0 as $t \to \infty$.

It is worth noting that the standard eigenvalue characterisation of asymptotic stability remains the same even for positive systems. In addition, we include the characterisation from Lyapunov stability theory.

Proposition 3.6. *The following statements are equivalent:*

1. *The positive system Σ_c^{auto} is asymptotically stable;*
2. *A is Hurwitz, i.e.* $\mathrm{Re}(\lambda) < 0$ *for all $\lambda \in \sigma(A)$;*
3. *There exists a $P \succ 0$ such that $A^T P + PA \prec 0$.*

Proposition 3.7. *The following statements are equivalent:*

1. *The positive system Σ_d^{auto} is asymptotically stable;*
2. *A is Schur, i.e. $|\lambda| < 1$ for all $\lambda \in \sigma(A)$;*
3. *There exists a $P \succ 0$ such that $A^T PA - P \prec 0$.*

To start with, we provide a quick test one can perform in order to check whether a positive system is *not* asymptotically stable.

Theorem 3.8. *Let Σ_c^{auto} and Σ_d^{auto} be positive.*

1. *If A is Hurwitz, then $a_{ii} \leq 0$ for all $i = 1, \ldots, n$.*
2. *If A is Schur, then $0 \leq a_{ii} \leq 1$ for all $i = 1, \ldots, n$.*

Proof. To prove (1), suppose that the ith diagonal element of A is positive. Additionally, from the positivity of $\Sigma_{\mathrm{c}}^{\mathrm{auto}}$, we know that $a_{ij} \geq 0$ for all $j \neq i$. Choose an initial condition $\mathbf{x}(0)$ such that $x_i(0) > 0$ and consider the ith element of $\dot{\mathbf{x}}(t)$:

$$\dot{x}_i(t) = \sum_{k=1}^{n} a_{ik}x_k(t) > a_{ii}x_i(t) > 0 \;\text{ for all }\; t \in \mathbb{R}_+.$$

Therefore, $\lim_{t\to\infty} x_i(t) = \infty$.

To prove (2), assume that $a_{ii} > 1$. Take $\mathbf{x}(0) = \mathbf{e}_i$ and using the fact that $A \geq 0$, we can show the following by induction:

$$\begin{aligned} x_i(1) &= \sum_{k=1}^{n} a_{ik} x_k(0) = a_{ii}, \\ x_i(2) &= \sum_{k=1}^{n} a_{ik} x_k(1) \geq a_{ii}^2, \\ &\vdots \\ x_i(t) &\geq a_{ii}^t, \end{aligned}$$

but because $a_{ii} > 1$, we know that the ith element diverges as $t \to \infty$. ■

Asymptotic stability of positive systems has been extensively studied with [2, Chapter 6] serving as a prime example – the authors provide 50 necessary and sufficient conditions (for continuous-time systems). Nevertheless, in this thesis, we focus only on a few, which have important consequences (see Subsection 3.2.2).

Theorem 3.9. *Let Σ_c^{auto} be positive. The following statements are equivalent:*

1. *A is Hurwitz;*
2. *A is invertible and $-A^{-1} \geq 0$;*
3. *There exists a $\boldsymbol{\xi} > 0$ such that $A\boldsymbol{\xi} < 0$;*
4. *There exists an $\boldsymbol{\eta} > 0$ such that $\boldsymbol{\eta}^T A < 0$;*
5. *There exists a diagonal matrix $P \succ 0$ such that $A^T P + PA \prec 0$.*

Remark. In other resources, e.g. [2, 4], asymptotic stability of a positive continuous-time system is studied under the name *nonsingular M-matrix* instead of Hurwitz Metzler matrix.

Theorem 3.10. *Let Σ_d^{auto} be positive. The following statements are equivalent:*

1. *A is Schur;*
2. *$I - A$ is invertible and $(I - A)^{-1} \geq 0$;*
3. *There exists a $\boldsymbol{\xi} > 0$ such that $A\boldsymbol{\xi} < \boldsymbol{\xi}$;*
4. *There exists an $\boldsymbol{\eta} > 0$ such that $\boldsymbol{\eta}^T A < \boldsymbol{\eta}^T$;*
5. *There exists a diagonal matrix $P \succ 0$ such that $A^T PA - P \prec 0$.*

We save the proofs of Theorems 3.9 and 3.10 for the next subsection.

3.2.1 Proofs of Theorems 3.9 and 3.10

A paramount result, on which positive systems theory strongly relies, is the Perron-Frobenius theorem.

Proposition 3.11. *Let A be a real square matrix.*

1. *If $A > 0$, then there exists an $\boldsymbol{x} > 0$ such that $A\boldsymbol{x} = \rho(A)\boldsymbol{x}$.*
2. *If $A \geq 0$, then there exists a nonzero $\boldsymbol{x} \geq 0$ such that $A\boldsymbol{x} = \rho(A)\boldsymbol{x}$.*

The first statement is part of the results on positive matrices, which Oskar Perron published in 1907 [8] and the second statement is a generalisation of his results to nonnegative matrices. Ferdinand G. Frobenius' contribution (which is not included in Proposition 3.11) comes in 1912 [9] when he adds an additional assumption on $A \geq 0$, so that we can guarantee the existence of a positive eigenvector, but it is out of the scope of this thesis. We refer the interested reader to [10, Chapter 8].

We begin with the proof of Theorem 3.9.

Proof of Theorem 3.9. We show that (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (5) $\Rightarrow$ (1).

(1 $\Rightarrow$ 2): The eigenvalues of a Hurwitz matrix are nonzero, so the inverse of A exists. Consider the equilibrium $\mathbf{x}_{\text{eq}}$ of the positive system $\dot{\mathbf{x}} = A\mathbf{x} + \bar{\mathbf{u}}$, where $\bar{\mathbf{u}} \geq 0$ is an arbitrary constant input. Then, $0 = A\mathbf{x}_{\text{eq}} + \bar{\mathbf{u}}$ implies that $\mathbf{x}_{\text{eq}} = -A^{-1}\bar{\mathbf{u}}$. The equilibrium is nonnegative due to the positivity of the system and because the input is arbitrary, we have that $-A^{-1} \geq 0$.

(2 $\Rightarrow$ 3): A Metzler matrix can be written as $A = B - cI$, where $c > 0$ is a scalar and, depending on A, either $B > 0$ or $B \geq 0$. If $B > 0$, by Proposition 3.11, there exists a $\boldsymbol{\xi} > 0$ such that $B\boldsymbol{\xi} = \rho(B)\boldsymbol{\xi}$. Then, $A\boldsymbol{\xi} = (B - cI)\boldsymbol{\xi} = (\rho(B) - c)\boldsymbol{\xi}$. After multiplying by $-A^{-1}$ on both sides, we have

$$(\rho(B) - c)(-A^{-1})\boldsymbol{\xi} = -\boldsymbol{\xi} < 0,$$

which is possible only if $\rho(B) - c < 0$, because $-A^{-1}\boldsymbol{\xi}$ is nonnegative and nonzero. This implies that $A\boldsymbol{\xi} < 0$.

If instead $B \geq 0$, by Proposition 3.11, there exists a nonzero $\mathbf{v} \geq 0$ such that $B\mathbf{v} = \rho(B)\mathbf{v}$. Similarly, we conclude that

$$(\rho(B) - c)(-A^{-1})\mathbf{v} = -\mathbf{v} \leq 0.$$

Because $-A^{-1}\mathbf{v}$ is nonnegative and nonzero, we know that $\rho(B) - c < 0$. Using the fact that eigenvalues are a continuous function of the matrix elements, we can increase the zero elements of B (and, consequently, those of A) by a small $\epsilon > 0$, the result of which we denote by B_ϵ (A_ϵ), such that $\rho(B_\epsilon) - c < 0$. Moreover, note that $A_\epsilon \geq A$. Then, by Proposition 3.11, there exists a $\boldsymbol{\xi} > 0$ such that

$$A\boldsymbol{\xi} \leq A_\epsilon\boldsymbol{\xi} = (B_\epsilon - cI)\boldsymbol{\xi} = (\rho(B_\epsilon) - c)\boldsymbol{\xi} < 0.$$

(3 $\Rightarrow$ 4): Let $A\boldsymbol{\xi} < 0$ for some $\boldsymbol{\xi} > 0$. We again express A as $A = B - cI$ and distinguish between the cases $B > 0$ and $B \geq 0$. Suppose that $B > 0$. By Proposition 3.11, there exists an $\boldsymbol{\eta} > 0$ such that $\boldsymbol{\eta}^T B = \boldsymbol{\eta}^T \rho(B)$. Because $A\boldsymbol{\xi} < 0$ implies that $\boldsymbol{\eta}^T A\boldsymbol{\xi} < 0$, we have

$$0 > \boldsymbol{\eta}^T A\boldsymbol{\xi} = \boldsymbol{\eta}^T (B - cI)\boldsymbol{\xi} = (\rho(B) - c)\boldsymbol{\eta}^T\boldsymbol{\xi}.$$

The vectors $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ are such that $\boldsymbol{\eta}^T\boldsymbol{\xi} > 0$, which means $\rho(B) - c < 0$. Subsequently,

$$\boldsymbol{\eta}^T A = \boldsymbol{\eta}^T (B - cI) = \boldsymbol{\eta}^T (\rho(B) - c) < 0.$$

Now, assume that $B \geq 0$. By Proposition 3.11, there exists a nonzero $\mathbf{w} \geq 0$ such that $\mathbf{w}^T B = \mathbf{w}^T \rho(B)$ and this time

$$0 > \mathbf{w}^T A\boldsymbol{\xi} = \mathbf{w}^T (B - cI)\boldsymbol{\xi} = (\rho(B) - c)\mathbf{w}^T\boldsymbol{\xi} \quad \text{and} \quad \mathbf{w}^T\boldsymbol{\xi} > 0 \quad \text{imply that} \quad \rho(B) - c < 0.$$

Note that the inequalities are strict, because $\mathbf{w}$ is nonzero. After using the same ϵ-perturbation argument as in the proof of (2 $\Rightarrow$ 3), we conclude that there exists an $\boldsymbol{\eta} > 0$ such that

$$\boldsymbol{\eta}^T A \leq \boldsymbol{\eta}^T A_\epsilon = \boldsymbol{\eta}^T (B_\epsilon - cI) = \boldsymbol{\eta}^T (\rho(B_\epsilon) - c) < 0.$$

(4 $\Rightarrow$ 5): The implication (4 $\Rightarrow$ 3) can easily be shown by using the same ideas from the proof of (3 $\Rightarrow$ 4). Therefore, we can assume that there exist positive $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ such that $A\boldsymbol{\xi}$ and $\boldsymbol{\eta}^T A$ are negative.

Construct an $n \times n$ matrix C with elements $c_{ij} := -\eta_i a_{ij} \xi_j$. From the positivity of $\Sigma_{\text{c}}^{\text{auto}}$, we know that $a_{ij} \geq 0$ for $j \neq i$. Additionally, Theorem 3.8 ensures that $a_{ii} \leq 0$. Then, $c_{ij} \leq 0$ for $j \neq i$ and $c_{ii} \geq 0$. We can show that both C and C^T are diagonally dominant:

$$\begin{aligned} |c_{ii}| - \sum_{\substack{j=1 \\ j\neq i}}^{n} |c_{ij}| &= \sum_{j=1}^{n} c_{ij} = -\eta_i \sum_{j=1}^{n} a_{ij}\xi_j = -\eta_i (A\xi)_i > 0, \\ |c_{ii}| - \sum_{\substack{j=1 \\ j\neq i}}^{n} |c_{ji}| &= \sum_{j=1}^{n} c_{ji} = -\xi_i \sum_{j=1}^{n} \eta_j a_{ji} = -\xi_i (\eta^T A)_i > 0. \end{aligned} \tag{3.2}$$

Thus, $C + C^T$ is also diagonally dominant. Because $C + C^T$ is symmetric, it has real eigenvalues and, by the Gershgorin Circle Theorem [7, Fact 6.10.22], they are in the union of the n open intervals, centered around $2c_{ii}$ with length $2\sum_{j=1;\, j\neq i}^{n} |c_{ij}|$. Furthermore, all intervals are fully contained in the right half of the real line, due to the diagonal dominance and that $c_{ii} > 0$ (as $c_{ii} \geq 0$ and (3.2) hold). Hence, $C + C^T$ has positive eigenvalues. By [11, Proposition 4.21], for all nonzero $\mathbf{x} \in \mathbb{R}^n$,

$$\mathbf{x}^T(C + C^T)\mathbf{x} = 2 \cdot \mathbf{x}^T C\mathbf{x} \geq \lambda_{\min}(C + C^T)\|\mathbf{x}\|^2 > 0,$$

where $\lambda_{\min}(C + C^T) > 0$ denotes the smallest eigenvalue of $C + C^T$ and $\|\cdot\|$ is the Euclidean norm.

Consider the diagonal matrix P with elements $p_{ii} := \eta_i/\xi_i > 0$, $i = 1, \ldots, n$. Then, $P \succ 0$. It remains to check whether $A^T P + PA$ is negative definite:

$$\begin{aligned}
-\mathbf{x}^T PA\mathbf{x} = -\sum_{i,j=1}^{n} x_i p_{ii} a_{ij} x_j &= \sum_{i,j=1}^{n} x_i \frac{\eta_i}{\xi_i} \frac{c_{ij}}{\eta_i \xi_j} x_j \\
&= \sum_{i,j=1}^{n} \frac{x_i}{\xi_i} c_{ij} \frac{x_j}{\xi_j} = \bar{\mathbf{x}}^T C \bar{\mathbf{x}} \qquad (\bar{\mathbf{x}} \in \mathbb{R}^n,\ \bar{x}_i := x_i/\xi_i,\ i = 1, \ldots, n) \\
&\geq \frac{\lambda_{\min}(C + C^T)}{2} \|\bar{\mathbf{x}}\|^2 > 0
\end{aligned}$$

for all nonzero $\mathbf{x} \in \mathbb{R}^n$, which means that PA as well as $A^T P$ are negative definite. Therefore, $A^T P + PA \prec 0$.

(5 ⇒ 1): This follows from Proposition 3.6. ■

Next, we prove Theorem 3.10 in a similar fashion.

Proof of Theorem 3.10. The proof strategy is the same as the one for Theorem 3.9.

(1 ⇒ 2): If $\lambda \in \sigma(A)$, then $1 - \lambda \in \sigma(I - A)$. In addition, A is Schur, which means that $1 - \lambda \neq 0$ and, thus, $I - A$ is invertible. Consider a positive system with an arbitrary constant input $\bar{\mathbf{u}} \geq 0$ and its equilibrium $\mathbf{x}_{eq}$. Then, $\mathbf{x}_{eq} = A\mathbf{x}_{eq} + \bar{\mathbf{u}}$, leading to $\mathbf{x}_{eq} = (I - A)^{-1}\bar{\mathbf{u}}$. Because the system is positive, the equilibrium has to be nonnegative, so we have that $(I - A)^{-1} \geq 0$.

(2 ⇒ 3): On the one hand, if $A > 0$, by Proposition 3.11, there exists a $\boldsymbol{\xi} > 0$ such that $A\boldsymbol{\xi} = \rho(A)\boldsymbol{\xi}$. Then, $(I - A)\boldsymbol{\xi} = (1 - \rho(A))\boldsymbol{\xi}$ and after multiplying both sides by the inverse of $I - A$, we have

$$(1 - \rho(A))(I - A)^{-1}\boldsymbol{\xi} = \boldsymbol{\xi} > 0,$$

which implies that $\rho(A) < 1$, because $(I - A)^{-1}\boldsymbol{\xi} > 0$. Therefore, $A\boldsymbol{\xi} < \boldsymbol{\xi}$.

On the other hand, if $A \geq 0$, by Proposition 3.11, there exists a nonzero $\mathbf{v} \geq 0$ such that $A\mathbf{v} = \rho(A)\mathbf{v}$. Similarly,

$$(1 - \rho(A))(I - A)^{-1}\mathbf{v} = \mathbf{v} \geq 0,$$

from which we again derive that $\rho(A) < 1$. After increasing the zero elements of A by a small $\epsilon > 0$ such that $\rho(A_\epsilon) < 1$, we see that

$$A\boldsymbol{\xi} \leq A_\epsilon \boldsymbol{\xi} = \rho(A_\epsilon)\boldsymbol{\xi} < \boldsymbol{\xi},$$

where $\boldsymbol{\xi} > 0$ is the right Perron eigenvector of A_ϵ.

(3 ⇒ 4): Assume $A\boldsymbol{\xi} < \boldsymbol{\xi}$ for some $\boldsymbol{\xi} > 0$. Firstly, let $A > 0$. By Proposition 3.11, there exists an $\boldsymbol{\eta} > 0$ such that $\boldsymbol{\eta}^T A = \boldsymbol{\eta}^T \rho(A)$. After right multiplication by $\boldsymbol{\xi}$, we have $\rho(A)\boldsymbol{\eta}^T\boldsymbol{\xi} = \boldsymbol{\eta}^T A\boldsymbol{\xi} < \boldsymbol{\eta}^T\boldsymbol{\xi}$ implying that $\rho(A) < 1$ and, thus, $\boldsymbol{\eta}^T A < \boldsymbol{\eta}^T$.

Secondly, let $A \geq 0$. By Proposition 3.11, there exists a nonzero $\mathbf{w} \geq 0$ such that $\mathbf{w}^T A = \mathbf{w}^T \rho(A)$. As before, right multiplication by $\boldsymbol{\xi}$ leads to $\rho(A)\mathbf{w}^T\boldsymbol{\xi} = \mathbf{w}^T A\boldsymbol{\xi} < \mathbf{w}^T\boldsymbol{\xi}$ and we again conclude that

$\rho(A) < 1$. After applying the same ϵ-perturbation argument on A as in the proof of ($2 \Rightarrow 3$), we know that

$$\boldsymbol{\eta}^T A \leq \boldsymbol{\eta}^T A_\epsilon = \boldsymbol{\eta}^T \rho(A_\epsilon) < \boldsymbol{\eta}^T,$$

where $\boldsymbol{\eta} > 0$ is the left Perron eigenvector of A_ϵ.

($4 \Rightarrow 5$): By following the ideas from the proof of ($3 \Rightarrow 4$), the implication ($4 \Rightarrow 3$) becomes straightforward. Then, we know that there exist positive $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ such that $A\boldsymbol{\xi} < \boldsymbol{\xi}$ and $\boldsymbol{\eta}^T A < \boldsymbol{\eta}$.

Define $E \coloneqq \operatorname{diag}(\boldsymbol{\eta})$ and $X \coloneqq \operatorname{diag}(\boldsymbol{\xi})$. Construct an $n \times n$ matrix D with elements $d_{ij} \coloneqq \eta_i a_{ij} \xi_j$, i.e. $D = EAX$. The positivity of $\Sigma_{\mathrm{d}}^{\mathrm{auto}}$ ensures that $d_{ij} \geq 0$ for all $i, j = 1, \ldots, n$. Now, we show that the matrix

$$F \coloneqq \begin{bmatrix} EX & D^T \\ D & EX \end{bmatrix}$$

is diagonally dominant:

$$i = 1, \ldots, n\colon |f_{ii}| - \sum_{\substack{k=1 \\ k \neq i}}^{2n} |f_{ik}| = \eta_i \xi_i - \sum_{j=1}^{n} \eta_j a_{ji} \xi_i = (\boldsymbol{\eta}^T - \boldsymbol{\eta}^T A)_i \xi_i > 0,$$

$$i = n+1, \ldots, 2n\colon |f_{ii}| - \sum_{\substack{k=1 \\ k \neq i}}^{2n} |f_{ik}| = \eta_i \xi_i - \sum_{j=1}^{n} \eta_i a_{ij} \xi_j = \eta_i (\boldsymbol{\xi} - A\boldsymbol{\xi})_i > 0.$$

Arguing in the same way as in the proof of ($4 \Rightarrow 5$) from Theorem 3.9, we can deduce that L has positive eigenvalues, which implies that $L \succ 0$. Hence, by [12, Lemma A.2.1], the Schur complement [7, Definition 8.1.13] of L is also positive definite:

$$EX - D^T X^{-1} E^{-1} D \succ 0.$$

Let $\mathbf{z} \in \mathbb{R}^n$ be arbitrary and nonzero. Note that

$$\begin{aligned} \mathbf{z}^T \left(EX^{-1} - A^T EX^{-1} A \right) \mathbf{z} &= \mathbf{z}^T X^{-1} EXX^{-1} \mathbf{z} - \mathbf{z}^T X^{-1} X A^T EX^{-1} E^{-1} EAXX^{-1} \mathbf{z} \\ &= \bar{\mathbf{z}}^T EX \bar{\mathbf{z}} - \bar{\mathbf{z}}^T D^T X^{-1} E^{-1} D \bar{\mathbf{z}} \\ &= \bar{\mathbf{z}}^T \left(EX - D^T X^{-1} E^{-1} D \right) \bar{\mathbf{z}} > 0, \end{aligned}$$

where $\bar{\mathbf{z}} \in \mathbb{R}^n$ is defined as $\bar{\mathbf{z}} \coloneqq X^{-1}\mathbf{z}$. By taking $P \coloneqq EX^{-1} > 0$, which is diagonal and positive definite, we have shown that $A^T P A - P \prec 0$.

($5 \Rightarrow 1$): This follows from Proposition 3.7. ■

3.2.2 Discussion

When comparing the Lyapunov equation conditions for asymptotic stability of positive and unconstrained systems, we can see a very notable difference concerning the P matrix. Namely, for positive systems, P is *diagonal*. This phenomenon can also be seen in the results in Chapters 4 and 5. The diagonal structure of P offers a great advantage in the context of large-scale systems. Indeed, the number of parameters scales linearly with the number of states instead of quadratically. Furthermore, the proofs of Theorems 3.9 and 3.10 provide a way to construct it using the left and right Perron eigenvectors of A (or A_ϵ):

$$p_{ii} \coloneqq \frac{\eta_i}{\xi_i}.$$

Statements (3) and (4) from Theorems 3.9 and 3.10 are also of significant interest. They allow us to reformulate semidefinite programming (SDP) problems into linear programming (LP) ones. For example, we do this for the stability verification and (positive) stabilisation (see Chapter 4) problems. While the solutions from SDP and LP are theoretically the same, LP has a lower computational complexity compared to SDP.

Lastly, we comment on the geometric interpretation of statements (3), (4) and (5) from Theorems 3.9 and 3.10. Consider the following Lyapunov functions, which correspond to each statement:

$$
\begin{aligned}
V_3(\mathbf{x}) &= \max_{i=1,\dots,n} \frac{x_i}{\xi_i},\ \boldsymbol{\xi} > 0,\\
V_4(\mathbf{x}) &= \boldsymbol{\eta}^T \mathbf{x},\ \boldsymbol{\eta} > 0,\\
V_5(\mathbf{x}) &= \mathbf{x}^T P \mathbf{x},\ P \succ 0.
\end{aligned}
$$

Asymptotic stability of a positive system is equivalent to the existence of a Lyapunov function V_i $(i = 3, 4, 5)$ such that $\dot{V}_i(\mathbf{x}) < 0$ (in continuous time) or $V_i(A\mathbf{x}) - V_i(\mathbf{x}) < 0$ (in discrete time) for all nonzero $\mathbf{x}$. For the sake of illustration, we look at the two-dimensional case $n = 2$. Consider the level curves C_k $(k \in \mathbb{Z}_+)$, where V_i is constant, i.e.

$$
C_k \coloneqq \left\{ \mathbf{x} \in \mathbb{R}^2 \mid V_i(\mathbf{x}) = c_k \right\},
$$

which are shown in Figure 1. Then, the trajectories of an asymptotically stable system penetrate the level curves in a decreasing order $(0 = c_0 < c_1 < c_2 < \cdots)$, converging to the origin as $t \to \infty$.

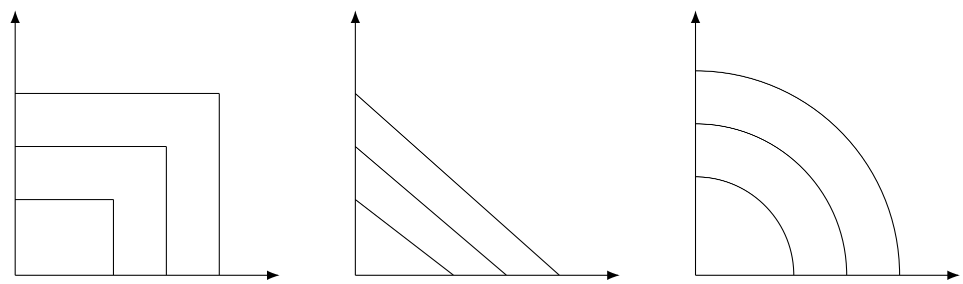

Figure 1: Level curves of the Lyapunov functions V_3, V_4 and V_5.

3.2.3 Examples

In this subsection, we look at specific numerical examples of (2.1) and (2.2) and check whether the systems are asymptotically stable.

Generally, the process of radioactive decay does not happen over the span of a few days or months. The half-life of Uranium-238, for instance, is 4.5 billion years. Therefore, we consider an unrealistic, yet illustrative example of Λ:

$$
\Lambda = \begin{bmatrix} -1 & & & & \\ 1 & -2 & & & \\ & 2 & -3 & & \\ & & \ddots & \ddots & \\ & & & 13 & -14 \end{bmatrix}.
$$

It can easily be verified that the following $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ satisfy statements (3) and (4) from Theorem 3.9:

$$
\boldsymbol{\xi} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix},\ \boldsymbol{\eta} = \begin{bmatrix} 14 \\ 13 \\ 12 \\ \vdots \\ 1 \end{bmatrix}.
$$

Additionally, instead of calculating a matrix P that satisfies statement (5) from Theorem 3.9 using

SDP, we can do so directly with $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ and verify that $A^TP + PA \prec 0$:

$$P = \begin{bmatrix} 14 & & & & \\ & 13 & & & \\ & & 12 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}.$$

For the Leslie model, we examine the first 10 age groups of a fish population [3, Chapter 13]:

s_0	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9
$6 \cdot 10^{-5}$	0.45	0.27	0.26	0.26	0.25	0.25	0.25	0.25	0.25

f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}
0	5000	11000	18000	24000	31000	34000	41000	45000	46000

While survival rates are quite low, the fish compensate with a high fertility rate that increases with age, which makes it an interesting case study. After checking the statements in Theorem 3.10 in MATLAB, we conclude that the fish go extinct:

$$\begin{aligned}\sigma(L) = \{&0.6485, 0.2047 + 0.2205i, 0.2047 - 0.2205i, 0.0463 + 0.3206i, 0.0463 - 0.3206i, \\ &-0.3384, -0.2663 + 0.1768i, -0.2663 - 0.1768i, -0.1398 + 0.2905i, -0.1398 - 0.2905i\},\end{aligned}$$

$$(I - L)^{-1} = \begin{bmatrix} 1.3632 & 0.8071 & 1.4745 & 2.2109 & 2.8411 & 3.5126 & 3.9081 & 4.5087 & 4.6212 & 3.7624 \\ 0.6134 & 1.3632 & 0.6635 & 0.9949 & 1.2785 & 1.5806 & 1.7586 & 2.0289 & 2.0795 & 1.6931 \\ 0.1656 & 0.3681 & 1.1792 & 0.2686 & 0.3452 & 0.4268 & 0.4748 & 0.5478 & 0.5615 & 0.4571 \\ 0.0431 & 0.0957 & 0.3066 & 1.0698 & 0.0898 & 0.1110 & 0.1235 & 0.1424 & 0.1460 & 0.1189 \\ 0.0112 & 0.0249 & 0.0797 & 0.2782 & 1.0233 & 0.0288 & 0.0321 & 0.0370 & 0.0380 & 0.0309 \\ 0.0028 & 0.0062 & 0.0199 & 0.0695 & 0.2558 & 1.0072 & 0.0080 & 0.0093 & 0.0095 & 0.0077 \\ 0.0007 & 0.0016 & 0.0050 & 0.0174 & 0.0640 & 0.2518 & 1.0020 & 0.0023 & 0.0024 & 0.0019 \\ 0.0002 & 0.0004 & 0.0012 & 0.0043 & 0.0160 & 0.0630 & 0.2505 & 1.0006 & 0.0006 & 0.0005 \\ 0.0000 & 0.0001 & 0.0003 & 0.0011 & 0.0040 & 0.0157 & 0.0626 & 0.2501 & 1.0001 & 0.0001 \\ 0.0000 & 0.0000 & 0.0001 & 0.0003 & 0.0010 & 0.0039 & 0.0157 & 0.0625 & 0.2500 & 1.0000 \end{bmatrix}$$

and by using YALMIP [13] with MOSEK [14], we obtain

$$\boldsymbol{\xi} = \begin{bmatrix} 2.9010 \\ 1.4054 \\ 0.4795 \\ 0.2247 \\ 0.1584 \\ 0.1396 \\ 0.1349 \\ 0.1337 \\ 0.1334 \\ 0.1334 \end{bmatrix}, \ \boldsymbol{\eta} = \begin{bmatrix} 0.2200 \\ 0.2667 \\ 0.3730 \\ 0.4915 \\ 0.5919 \\ 0.7001 \\ 0.7636 \\ 0.8590 \\ 0.8709 \\ 0.7073 \end{bmatrix}, \ P = \operatorname{diag}\left(\begin{bmatrix} 0.0333 \\ 0.0809 \\ 0.7782 \\ 1.1521 \\ 1.2575 \\ 1.3727 \\ 1.4295 \\ 1.5634 \\ 1.6409 \\ 1.5852 \end{bmatrix} \right).$$

Chapter 4

Kalman–Yakubovich–Popov Lemma

The Kalman–Yakubovich–Popov lemma is a fundamental result in modern control theory, which was first formulated and proven by Vladimir A. Yakubovich in 1962 [15]. The initial formulation was the equivalence between a strict frequency inequality and a matrix inequality. The result is extended to the nonstrict case by Rudolf E. Kalman in 1963 [16]. However, until then, Yakubovich and Kalman only consider one-dimensional inputs. The constraint on the control dimensionality is removed by Vasile M. Popov in 1964 [17]. We refer the interested reader to a historical essay about the KYP lemma [18].

Ultimately, the KYP lemma creates a bridge between the frequency domain and state space. Later, this connection is extended to positive systems in both continuous and discrete time by Anders Rantzer, originally, in 2012 [19] and in the later iterations [20, 21]. The results are presented below.

Theorem 4.1. *Assume A is a Hurwitz Metzler matrix, $B \geq 0$ and the pair (A, B) is controllable. Let $M \in \mathbb{S}^{n+m}$ be a matrix whose off-diagonal and first n diagonal elements are nonnegative. The following statements are equivalent:*

1. *For all $\omega \in \mathbb{R}$,*
$$\begin{bmatrix} (i\omega I - A)^{-1}B \\ I \end{bmatrix}^* M \begin{bmatrix} (i\omega I - A)^{-1}B \\ I \end{bmatrix} \preceq 0;$$

2. *The following inequality holds:*
$$\begin{bmatrix} -A^{-1}B \\ I \end{bmatrix}^T M \begin{bmatrix} -A^{-1}B \\ I \end{bmatrix} \preceq 0;$$

3. *There exists a diagonal $P \succeq 0$ such that*
$$M + \begin{bmatrix} A^T P + PA & PB \\ B^T P & 0 \end{bmatrix} \preceq 0.$$

If all inequalities in (1), (2) and (3) are replaced with strict ones, then the statements remain equivalent even without the controllability assumption.

Theorem 4.2. *Assume $A \geq 0$ is Schur, $B \geq 0$ and the pair (A, B) is controllable. Let $M \in \mathbb{S}^{n+m}$ be a matrix whose off-diagonal and first n diagonal elements are nonnegative. The following statements are equivalent:*

1. *For all $\omega \in \mathbb{R}$,*
$$\begin{bmatrix} (e^{i\omega} I - A)^{-1}B \\ I \end{bmatrix}^* M \begin{bmatrix} (e^{i\omega} I - A)^{-1}B \\ I \end{bmatrix} \preceq 0;$$

2. The following inequality holds:

$$\begin{bmatrix} (I-A)^{-1}B \\ I \end{bmatrix}^T M \begin{bmatrix} (I-A)^{-1}B \\ I \end{bmatrix} \preceq 0;$$

3. There exists a diagonal $P \succeq 0$ such that

$$M + \begin{bmatrix} A^TPA - P & A^TPB \\ B^TPA & B^TPB \end{bmatrix} \preceq 0.$$

If all inequalities in (1), (2) and (3) are replaced with strict ones, then the statements remain equivalent even without the controllability assumption.

Firstly, we note the importance of the controllability assumption for the nonstrict case. Consider the following counterexample for Theorem 4.1:

$$A = -1,\ B = 0,\ M = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Then,

$$\begin{bmatrix} -A^{-1}B \\ I \end{bmatrix}^T M \begin{bmatrix} -A^{-1}B \\ I \end{bmatrix} = 0 \preceq 0,$$

but the eigenvalues of

$$M + \begin{bmatrix} A^TP + PA & PB \\ B^TP & 0 \end{bmatrix} = \begin{bmatrix} -2p_{11} & 1 \\ 1 & 0 \end{bmatrix}$$

are $-p_{11} \pm \sqrt{p_{11}^2 + 1}$, so statement (3) cannot be satisfied for any $p_{11} \geq 0$.

Secondly, we mention two differences to the version for unconstrained systems. Similarly to statements (5) from Theorems 3.9 and 3.10, here, we require the P matrix to be diagonal. Additionally, while we still have a frequency inequality and a matrix inequality, in the case of positive systems, it is enough to verify whether the frequency inequality holds only for $\omega = 0$.

4.1 Preliminaries

In this section, we provide the tools we use to prove the KYP lemma for positive systems. The following is a result from [21, Theorem 4].

Proposition 4.3. *Let $M \in \mathbb{S}^n$ and $N \in \mathbb{R}^{n\times n}$ be Metzler. Then,*

$$\max\left\{\boldsymbol{z}^T M \boldsymbol{z} \mid \boldsymbol{z} \geq 0,\ N\boldsymbol{z} \geq 0,\ \boldsymbol{z}^T\boldsymbol{z} \leq 1\right\} = \max\left\{\operatorname{tr}(MZ) \mid Z \succeq 0,\ \operatorname{diag}(NZ) \geq 0,\ \operatorname{tr}(Z) \leq 1\right\}. \quad (4.1)$$

Moreover, (4.1) holds if the constraint $Z \succeq 0$ on the right-hand side of (4.1) is relaxed to $Z \in \mathbb{P}$, where

$$\mathbb{P} := \left\{Z \in \mathbb{S}^n \mid z_{ii} \geq 0,\ z_{ij}^2 \leq z_{ii}z_{jj} \text{ for all } i,j = 1,\dots,n\right\}.$$

Proof. If we choose $Z = \mathbf{z}\mathbf{z}^T$, where $\mathbf{z}$ satisfies the constraints on the left-hand side of (4.1), then we can show that Z satisfies the constraints on the right-hand side:

$$\begin{aligned} &\mathbf{x}^T Z\mathbf{x} = (\mathbf{z}^T\mathbf{x})^T\mathbf{z}^T\mathbf{x} \geq 0 \text{ for all nonzero } \mathbf{x} \in \mathbb{R}^n; \\ &\mathbf{z} \geq 0 \text{ and } N\mathbf{z} \geq 0 \text{ imply that } \operatorname{diag}(N\mathbf{z}\mathbf{z}^T) = \operatorname{diag}(NZ) \geq 0; \\ &\operatorname{tr}(Z) = \operatorname{tr}(\mathbf{z}\mathbf{z}^T) = \mathbf{z}^T\mathbf{z} \leq 1. \end{aligned}$$

Hence, the right-hand side of (4.1) is greater than or equal to the left-hand side.

Suppose $Z \succeq 0$, i.e. $\mathbf{x}^T Z \mathbf{x} \geq 0$ for all nonzero $\mathbf{x} \in \mathbb{R}^n$. Note that $Z \in \mathbb{P}$. Indeed, for all $i, j = 1, \ldots, n$ and $v, w \in \mathbb{R}$, we have

$$\begin{aligned}
\mathbf{x} = \mathbf{e}_i &\Rightarrow z_{ii} \geq 0; \\
\mathbf{x} = v\mathbf{e}_i + w\mathbf{e}_j &\Rightarrow \begin{bmatrix} v \\ w \end{bmatrix}^T \begin{bmatrix} z_{ii} & z_{ij} \\ z_{ji} & z_{jj} \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} \geq 0 \Rightarrow \begin{bmatrix} z_{ii} & z_{ij} \\ z_{ji} & z_{jj} \end{bmatrix} \succeq 0 \\
&\Rightarrow \det\left(\begin{bmatrix} z_{ii} & z_{ij} \\ z_{ji} & z_{jj} \end{bmatrix} \right) = z_{ii} z_{jj} - z_{ij}^2 \geq 0 \Rightarrow z_{ij}^2 \leq z_{ii} z_{jj} \Rightarrow z_{ij} \leq \sqrt{z_{ii} z_{jj}}.
\end{aligned}$$

Select $\mathbf{z} := \begin{bmatrix} \sqrt{z_{11}} & \cdots & \sqrt{z_{nn}} \end{bmatrix}^T \geq 0$. Then, $\mathbf{z}\mathbf{z}^T \geq Z$, though, they have the same diagonal elements. In addition, $\mathbf{z}^T \mathbf{z} = \operatorname{tr}(Z) \leq 1$. Using the fact that the off-diagonal elements of M are nonnegative, we see that

$$\mathbf{z}^T M \mathbf{z} = \sum_{i,j=1}^{n} \sqrt{z_{ii}} m_{ij} \sqrt{z_{jj}} \geq \sum_{i,j=1}^{n} m_{ij} z_{ij} = \operatorname{tr}(MZ).$$

Consider the (k,k)th element of $N\mathbf{z}\mathbf{z}^T$. Because N is Metzler and $\operatorname{diag}(NZ) \geq 0$, we see that

$$(N\mathbf{z}\mathbf{z}^T)_{kk} = \sum_{j=1}^{n} n_{kj} \sqrt{z_{jj}} \sqrt{z_{kk}} \geq \sum_{j=1}^{n} n_{kj} z_{jk} = (NZ)_{kk} \geq 0.$$

Due to the fact that $\sqrt{z_{kk}} \geq 0$, for all $k = 1, \ldots, n$, we have $\sum_{j=1}^{n} n_{kj} \sqrt{z_{jj}} = N_{k\bullet}\mathbf{z} \geq 0$ or, equivalently, $N\mathbf{z} \geq 0$. Therefore, the left-hand side of (4.1) is at least as big as the right-hand side. We conclude that both sides of (4.1) are equal. ■

Next, we briefly present some duality theory for linear matrix inequalities (LMIs) from [22]. Denote by $\mathcal{S}^n$ the space of $n \times n$ Hermitian matrices and by $\mathcal{S}$ the space of block diagonal Hermitian matrices, i.e. $\mathcal{S} := \mathcal{S}^{n_1} \times \cdots \times \mathcal{S}^{n_L}$, with inner product

$$\langle \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_L \end{bmatrix}, \begin{bmatrix} B_1 & & \\ & \ddots & \\ & & B_L \end{bmatrix} \rangle_{\mathcal{S}} = \sum_{k=1}^{L} \operatorname{tr}(A_k B_k).$$

Consider the linear mapping $\mathcal{A} : \mathcal{V} \to \mathcal{S}$, where $\mathcal{V}$ is a finite-dimensional vector space with an inner product $\langle \cdot, \cdot \rangle_{\mathcal{V}}$. The adjoint mapping $\mathcal{A}^{\text{adj}} : \mathcal{S} \to \mathcal{V}$ of $\mathcal{A}$ is such that $\langle \mathcal{A}(\mathbf{x}),\ Z \rangle_{\mathcal{S}} = \langle \mathbf{x},\ \mathcal{A}^{\text{adj}}(Z) \rangle_{\mathcal{V}}$ for all $\mathbf{x} \in \mathcal{V}$ and $Z \in \mathcal{S}$. We make use of the following *theorems of alternatives*.

Proposition 4.4. *Let $A_0 \in \mathcal{S}$. Exactly one of the following statements is true:*

1. *There exists an $x \in \mathcal{V}$ such that $\mathcal{A}(x) + A_0 \succ 0$;*
2. *There exists a nonzero positive semidefinite $Z \in \mathcal{S}$ such that $\mathcal{A}^{adj}(Z) = 0$ and $\langle A_0, Z \rangle_{\mathcal{S}} \leq 0$.*

Proposition 4.5. *Let $A_0 \in \mathcal{S}$. At most one of the following statements is true:*

1. *There exists an $x \in \mathcal{V}$ such that $\mathcal{A}(x) + A_0 \succeq 0$;*
2. *There exists a positive semidefinite $Z \in \mathcal{S}$ such that $\mathcal{A}^{adj}(Z) = 0$ and $\langle A_0, Z \rangle_{\mathcal{S}} < 0$.*

Moreover, if $A_0 = \mathcal{A}(x_0)$ for some $x_0 \in \mathcal{V}$ or if there exists no $x \in \mathcal{V}$ such that $\mathcal{A}(x)$ is nonzero and positive semidefinite, then exactly one of the two statements is true.

The statements in Proposition 4.4 are called *strong alternatives*, because exactly one of them is true, whereas the ones in Proposition 4.5 are called *weak alternatives*. The additional assumptions under which weak alternatives become strong alternatives, are called *constraint qualifications*.

4.2 Proofs of Theorems 4.1 and 4.2

Now, we are ready to prove the KYP lemma for positive systems.

Proof of Theorem 4.1. Introduce two additional statements:

5. The inequality
$$\begin{bmatrix}\mathbf{x}\\ \mathbf{u}\end{bmatrix}^T M \begin{bmatrix}\mathbf{x}\\ \mathbf{u}\end{bmatrix} \leq 0$$
holds for all $\mathbf{x}, \mathbf{u} \geq 0$ satisfying $A\mathbf{x} + B\mathbf{u} \geq 0$;
6. The inequality $\mathrm{tr}(MZ) \leq 0$ holds for all $Z \succeq 0$ satisfying
$$\mathrm{diag}\left(\begin{bmatrix}A & B\end{bmatrix} Z \begin{bmatrix}I & 0\end{bmatrix}^T\right) \geq 0.$$

We show that (3) $\Rightarrow$ (1) $\Rightarrow$ (2) $\Rightarrow$ (5) $\Rightarrow$ (6) $\Rightarrow$ (3).

(3 $\Rightarrow$ 1): Let $\omega \in \mathbb{R}$ and $\mathbf{u} \in \mathbb{C}^m$ and define $\mathbf{x} := (i\omega I - A)^{-1}B\mathbf{u}$. We obtain the desired inequality as follows:

$$\begin{aligned}
\begin{bmatrix}\mathbf{x}\\ \mathbf{u}\end{bmatrix}^* \left(M + \begin{bmatrix}A^TP + PA & PB\\ B^TP & 0\end{bmatrix}\right)\begin{bmatrix}\mathbf{x}\\ \mathbf{u}\end{bmatrix} &= \begin{bmatrix}\mathbf{x}\\ \mathbf{u}\end{bmatrix}^* M \begin{bmatrix}\mathbf{x}\\ \mathbf{u}\end{bmatrix} + \begin{bmatrix}\mathbf{x}\\ \mathbf{u}\end{bmatrix}^* \begin{bmatrix}I & 0\\ A & B\end{bmatrix}^T \begin{bmatrix}0 & P\\ P & 0\end{bmatrix}\begin{bmatrix}I & 0\\ A & B\end{bmatrix}\begin{bmatrix}\mathbf{x}\\ \mathbf{u}\end{bmatrix}\\
&= \begin{bmatrix}\mathbf{x}\\ \mathbf{u}\end{bmatrix}^* M \begin{bmatrix}\mathbf{x}\\ \mathbf{u}\end{bmatrix}\\
&= \mathbf{u}^* \begin{bmatrix}(i\omega I - A)^{-1}B\\ I\end{bmatrix}^* M \begin{bmatrix}(i\omega I - A)^{-1}B\\ I\end{bmatrix}\mathbf{u} \leq 0.
\end{aligned}$$

(1 $\Rightarrow$ 2): Take $\omega = 0$.

(2 $\Rightarrow$ 5): By Theorem 3.9, we know that $-A^{-1} \geq 0$. So if $A\mathbf{x} + B\mathbf{u} \geq 0$, where $\mathbf{x}, \mathbf{u} \geq 0$, then $\mathbf{x} \leq -A^{-1}B\mathbf{u}$. Due to the structure of M, we have

$$\begin{bmatrix}\mathbf{x}\\ \mathbf{u}\end{bmatrix}^T M \begin{bmatrix}\mathbf{x}\\ \mathbf{u}\end{bmatrix} \leq \mathbf{u}^T \begin{bmatrix}-A^{-1}B\\ I\end{bmatrix}^T M \begin{bmatrix}-A^{-1}B\\ I\end{bmatrix}\mathbf{u} \leq 0.$$

(5 $\Rightarrow$ 6): Define

$$N := \begin{bmatrix}A & B\\ 0 & 0\end{bmatrix}, \; \mathbf{z} := \begin{bmatrix}\mathbf{x}\\ \mathbf{u}\end{bmatrix}$$

and note that N is Metzler. If $\mathbf{x}, \mathbf{u} \geq 0$ are such that $A\mathbf{x} + B\mathbf{u} \geq 0$, then $\mathbf{z} \geq 0$ and $N\mathbf{z} \geq 0$. By assumption, we know that

$$\max\left\{\mathbf{z}^T M\mathbf{z} \mid \mathbf{z} \geq 0,\; N\mathbf{z} \geq 0\right\} \leq 0.$$

We can scale $\mathbf{z}$ such that $\mathbf{z}^T\mathbf{z} \leq 1$ and, consequently, obtain

$$\max\left\{\mathbf{z}^T M\mathbf{z} \mid \mathbf{z} \geq 0,\; N\mathbf{z} \geq 0,\; \mathbf{z}^T\mathbf{z} \leq 1\right\} \leq 0.$$

Hence, by Proposition 4.3,

$$\max\left\{\mathrm{tr}(MZ) \mid Z \succeq 0,\; \mathrm{diag}(NZ) \geq 0,\; \mathrm{tr}(Z) \leq 1\right\} \leq 0,$$

where

$$\mathrm{diag}(NZ) = \mathrm{diag}\left(\begin{bmatrix}A & B\\ 0 & 0\end{bmatrix} Z\right) = \mathrm{diag}\left(\begin{bmatrix}A & B\end{bmatrix} Z \begin{bmatrix}I & 0\end{bmatrix}^T\right) \geq 0.$$

Similarly, because Z is subject to scaling, we can conclude that $\mathrm{tr}(MZ) \leq 0$ for all $Z \succeq 0$ satisfying $\mathrm{diag}\left(\begin{bmatrix}A & B\end{bmatrix} Z \begin{bmatrix}I & 0\end{bmatrix}^T\right) \geq 0$.

(6 ⇒ 3): Consider the spaces $\mathcal{V} := \mathbb{R}^n$ and

$$\mathcal{S} := \left\{ Q := \left[\begin{array}{cc:c} Q_{11} & Q_{12} & \\ Q_{12}^T & Q_{22} & \\ \hdashline & & Q_{33} \end{array}\right] \mid Q_{11} \in \mathbb{S}^n,\ Q_{22} \in \mathbb{S}^m,\ Q_{12} \in \mathbb{R}^{n\times m},\ Q_{33} := \operatorname{diag}(\mathbf{q}),\ \mathbf{q} \in \mathbb{R}^n \right\}$$

with inner products $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{V}} = \mathbf{x}^T\mathbf{y}$ and $\langle A, B\rangle_{\mathcal{S}} = \operatorname{tr}(AB)$. We can rephrase statement (3) as follows: there exists a $\mathbf{p} \in \mathcal{V}$ such that

$$\mathcal{A}(\mathbf{p}) + A_0 := \left[\begin{array}{cc:c} -A^TP - PA & -PB & \\ -B^TP & 0 & \\ \hdashline & & P \end{array}\right] + \left[\begin{array}{c:c} -M & \\ \hdashline & 0 \end{array}\right] \succeq 0, \tag{4.2.1}$$

where $P := \operatorname{diag}(\mathbf{p})$. Note that

$$\begin{aligned}
&\left\langle \left[\begin{array}{cc:c} -A^TP - PA & -PB & \\ -B^TP & 0 & \\ \hdashline & & P \end{array}\right], \left[\begin{array}{cc:c} Q_{11} & Q_{12} & \\ Q_{12}^T & Q_{22} & \\ \hdashline & & Q_{33} \end{array}\right] \right\rangle_{\mathcal{S}} \\
&= \operatorname{tr}\left(\begin{bmatrix} -A^TP - PA & -PB \\ -B^TP & 0 \end{bmatrix} \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{bmatrix} \right) + \mathbf{p}^T\mathbf{q} \\
&= -\operatorname{tr}(A^TPQ_{11}) - \operatorname{tr}(PAQ_{11}) - \operatorname{tr}(PBQ_{12}^T) - \operatorname{tr}(B^TPQ_{12}) + \mathbf{p}^T\mathbf{q} \\
&= -\mathbf{p}^T(\operatorname{diag}(Q_{11}A^T) + \operatorname{diag}(AQ_{11}) + \operatorname{diag}(BQ_{12}^T) + \operatorname{diag}(Q_{12}B^T)) + \mathbf{p}^T\mathbf{q} \\
&= -2 \cdot \mathbf{p}^T \operatorname{diag}(AQ_{11} + BQ_{12}^T) + \mathbf{p}^T\mathbf{q} \\
&= \langle \mathbf{p}, -2 \cdot \operatorname{diag}(AQ_{11} + BQ_{12}^T) + \mathbf{q} \rangle_{\mathcal{V}}.
\end{aligned}$$

Therefore, the adjoint mapping of $\mathcal{A}$ is $\mathcal{A}^{\text{adj}}(Q) = -2 \cdot \operatorname{diag}(AQ_{11} + BQ_{12}^T) + \mathbf{q}$. Additionally, if $Q \succeq 0$, then

$$\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{bmatrix} \succeq 0 \text{ and } \mathbf{q} \geq 0,$$

so $\mathcal{A}^{\text{adj}}(Q) = 0$ implies that

$$\operatorname{diag}(AQ_{11} + BQ_{12}^T) = \operatorname{diag}\left(\begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{bmatrix} \begin{bmatrix} I & 0 \end{bmatrix}^T \right) \geq 0.$$

Assume that there exists a $\mathbf{p} \in \mathcal{V}$ such that $\mathcal{A}(\mathbf{p})$ is nonzero and positive semidefinite or, equivalently, $P \succeq 0$ and

$$\begin{bmatrix} -A^TP - PA & -PB \\ -B^TP & 0 \end{bmatrix} \succeq 0$$

are nonzero. By [7, Corollary 10.2.3], we must have that $-B^TP = 0$ and, consequently, $-PB = 0$. Thus, $A^TP + PA \preceq 0$. From the assumption that (A, B) is controllable, by [11, Theorem 9.2], we can choose a matrix K such that all eigenvalues of $A + BK$ are distinct and have positive real parts. This means its eigenvectors are linearly independent. However, if all of them are in the kernel of P, then, by the rank-nullity theorem [7, Corollary 3.6.5], P has to be the zero matrix. Hence, there exists a $\mathbf{v}$ such that $P\mathbf{v} \neq 0$ and $(A + BK)\mathbf{v} = \lambda\mathbf{v}$. Then,

$$\begin{aligned}
\mathbf{v}^* \left(A^TP + PA\right) \mathbf{v} = \mathbf{v}^* \left((A + BK)^TP + P(A + BK)\right) \mathbf{v} \leq 0, \\
2 \cdot \operatorname{Re}(\lambda)(\mathbf{v}^*P\mathbf{v}) \leq 0.
\end{aligned}$$

From the positive semidefiniteness of P, we conclude that $\operatorname{Re}(\lambda) \leq 0$, which is a contradiction.

Finally, by Theorem 4.5, exactly one of the following two statements is true:

1. There exists a $\mathbf{p} \in \mathcal{V}$ such that (4.2.1) holds;
2. There exists a $\hat{Q} \succeq 0$ such that $\operatorname{diag}\left(\begin{bmatrix} A & B \end{bmatrix} \hat{Q} \begin{bmatrix} I & 0 \end{bmatrix}^T\right) \geq 0$ and $\langle -M, \hat{Q} \rangle_{\mathcal{S}} = -\operatorname{tr}(M\hat{Q}) < 0$.

By assuming (6), clearly, we see that the second statement is false.

In case of strict inequalities, proving the equivalence between (1), (2) and (3) is a matter of replacing nonstrict inequalities with strict ones and, instead of Proposition 4.5, we invoke Proposition 4.4 and skip the arguments involving the controllability of (A, B). ■

Theorem 4.2 can be proven in a similar way, but we show a shorter proof, which utilises the already proven Theorem 4.1.

Proof of Theorem 4.2. We can use the Cayley transform $(1 + i\omega)/(1 - i\omega)$ to map the real numbers to the unit disk, instead of $e^{i\omega}$. By following the ideas in the proof of $(3 \Rightarrow 1)$ from Theorem 4.1, we see that statement (1) is equivalent to

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix}^* M \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} \leq 0, \tag{4.2.2}$$

where $(\mathbf{x}, \mathbf{u}, \omega)$ are such that $\mathbf{x} = (((1 + i\omega)/(1 - i\omega))I - A)^{-1} B\mathbf{u}$. Define

$$\begin{aligned}
\widehat{A} &\coloneqq (A - I)(A + I)^{-1}, \\
\widehat{B} &\coloneqq 2 \cdot (A + I)^{-1} B, \\
\widehat{\mathbf{x}} &\coloneqq \mathbf{x} + A\mathbf{x} + B\mathbf{u}, \\
S &\coloneqq \begin{bmatrix} (A + I)^{-1} & -(A + I)^{-1} B \\ 0 & I \end{bmatrix}, \\
\widehat{M} &\coloneqq S^T M S.
\end{aligned}$$

Then, we can rewrite (4.2.2) as

$$\begin{bmatrix} \widehat{\mathbf{x}} \\ \mathbf{u} \end{bmatrix}^* \widehat{M} \begin{bmatrix} \widehat{\mathbf{x}} \\ \mathbf{u} \end{bmatrix} \leq 0,$$

where $(\widehat{\mathbf{x}}, \mathbf{u}, \omega)$ are such that $\widehat{\mathbf{x}} = (i\omega I - \widehat{A})^{-1} \widehat{B}\mathbf{u}$, and, thus, we return to the setting of Theorem 4.1. After taking $\omega = 0$, we obtain the inequality from statement (2). By Theorem 4.1, we know that there exists a diagonal $Q \succeq 0$ such that

$$\widehat{M} + \begin{bmatrix} \widehat{A}^T Q + Q\widehat{A} & Q\widehat{B} \\ \widehat{B}^T Q & 0 \end{bmatrix} \preceq 0.$$

Let $P \coloneqq 2Q$ and note that

$$\begin{aligned}
&\begin{bmatrix} A + I & B \\ 0 & I \end{bmatrix}^T \left(\widehat{M} + \begin{bmatrix} \widehat{A}^T Q + Q\widehat{A} & Q\widehat{B} \\ \widehat{B}^T Q & 0 \end{bmatrix} \right) \begin{bmatrix} A + I & B \\ 0 & I \end{bmatrix} \\
&= M + 2 \cdot \begin{bmatrix} A^T Q A - Q & A^T Q B \\ B^T Q A & B^T Q B \end{bmatrix} = M + \begin{bmatrix} A^T P A - P & A^T P B \\ B^T P A & B^T P B \end{bmatrix} \preceq 0.
\end{aligned}$$

Proving the version of Theorem 4.2 with strict inequalities is again a matter of replacing nonstrict inequalities with strict ones. ■

4.3 A Numerical Example

Consider an example of a continuous-time system provided in [23, Section V][1]:

$$A = \begin{bmatrix} -0.3329 & 0.0800 & 0 & 0.0700 \\ 0.0800 & -0.2800 & 0.1000 & 0 \\ 0 & 0.1000 & -0.9650 & 0.0900 \\ 0.0700 & 0 & 0.0900 & -0.2000 \end{bmatrix}, \; B = I,$$

$$M = \begin{bmatrix} 0.2400 & 0.3613 & 0.2197 & 0.3400 & & & & \\ 0.3613 & 1.5323 & 0.8102 & 1.4203 & & & & \\ 0.2197 & 0.8102 & 0.7826 & 0.8622 & & & & \\ 0.3400 & 1.4203 & 0.8622 & 2.1021 & & & & \\ & & & & -151.7824 & & & \\ & & & & & -151.7824 & & \\ & & & & & & -151.7824 & \\ & & & & & & & -151.7824 \end{bmatrix}.$$

We use MATLAB to calculate the eigenvalues of

$$\begin{bmatrix} -A^{-1}B \\ I \end{bmatrix}^T M \begin{bmatrix} -A^{-1}B \\ I \end{bmatrix},$$

which are $\{-151.2841,\ -151.6803,\ -151.7716,\ -151.7603\}$. After verifying that A is Hurwitz, by Theorem 4.1, we know that

$$\begin{bmatrix} (i\omega I - A)^{-1}B \\ I \end{bmatrix}^* M \begin{bmatrix} (i\omega I - A)^{-1}B \\ I \end{bmatrix} \prec 0$$

for all $\omega \in \mathbb{R}$ and that there exists a diagonal $P \succ 0$ such that

$$M + \begin{bmatrix} A^T P + PA & PB \\ B^T P & 0 \end{bmatrix} \prec 0.$$

We find P by using YALMIP with MOSEK:

$$P = \mathrm{diag}\left(\begin{bmatrix} 18.7093 \\ 25.4319 \\ 27.6458 \\ 22.0076 \end{bmatrix} \right).$$

[1] In this paper, the authors apply the KYP lemma for positive systems to the problem of synthesising stabilising controllers with guaranteed performance, i.e. the $\mathcal{H}_\infty$ problem.

Chapter 5

Positive Stabilisation

In general, unstable behaviour of physical systems is undesirable. This chapter deals with the problem of stabilisation by state feedback, i.e. we seek to design a controller $\mathbf{u}(t) = K\mathbf{x}(t)$, which makes the continuous- or discrete-time closed-loop system

$$\Sigma_{\mathrm{c}}^{\mathrm{cl}}: \begin{cases} \dot{\mathbf{x}}(t) = (A+BK)\mathbf{x}(t), \\ \mathbf{y}(t) = (C+DK)\mathbf{x}(t), \end{cases} \qquad \Sigma_{\mathrm{d}}^{\mathrm{cl}}: \begin{cases} \mathbf{x}(t+1) = (A+BK)\mathbf{x}(t), \\ \mathbf{y}(t) = (C+DK)\mathbf{x}(t), \end{cases}$$

asymptotically stable. However, the standard stabilisation problem does not guarantee that the resulting system remains positive. Hence, in the context of positive systems, we discuss the problem of *positive stabilisation* – we require that $A+BK$ is Hurwitz Metzler or nonnegative Schur as well as $C + DK \geq 0$. In Section 5.1, we provide a method, originally formulated in [24], to compute such a K by SDP. In Section 5.2, based on [25], we reduce the computational complexity by transforming it into an LP problem instead. Finally, we would like to stress the following: (1) we do *not* assume that the original system is positive (see Section 5.3) and (2) the results in this chapter provide *necessary and sufficient* conditions for the existence of K.

5.1 By Semidefinite Programming

The next two results provide a necessary and sufficient LMI condition for the existence of a state feedback controller. The basis for the first result is statement (5) from Theorem 3.9.

Theorem 5.1. *Given the system Σ_c, there exists a feedback gain K, which makes the closed-loop system Σ_c^{cl} positive and asymptotically stable if and only if there exist an $n \times n$ diagonal matrix $Q \succ 0$ and a matrix $R \in \mathbb{R}^{m\times n}$ such that $AQ + BR$ is Metzler, $CQ + DR \geq 0$ and*

$$(AQ + BR)^T + AQ + BR \prec 0.$$

Moreover, a feedback gain, which achieves the desired objective, is given by $K = RQ^{-1}$.

Proof. **(⇒):** Suppose there exists a K such that $A + BK$ is Hurwitz Metzler and $C + DK \geq 0$. Firstly, by Theorem 3.9, there exists a diagonal $P \succ 0$ such that

$$(A + BK)^T P + P(A + BK) \prec 0.$$

After applying the following congruence transformation, we obtain

$$\begin{aligned} P^{-1}\left((A+BK)^T P + P(A+BK)\right)P^{-1} &= ((A+BK)P^{-1})^T + (A+BK)P^{-1} \\ &= (AQ+BR)^T + AQ + BR \prec 0, \end{aligned}$$

where we define $Q \coloneqq P^{-1} \succ 0$ and $R \coloneqq KQ$, i.e. $K = RQ^{-1}$.

Secondly, because $Q > 0$, note that if $A + BK = A + BRQ^{-1}$ is Metzler and $C + DK = C + DRQ^{-1} \geq 0$, then also $AQ + BR$ is Metzler and $CQ + DR \geq 0$.

($\Leftarrow$): Let $K = RQ^{-1}$. Because $Q \succ 0$, consequently, $Q > 0$ and $Q^{-1} > 0$. Thus, if $AQ + BR$ is Metzler and $CQ + DR \geq 0$, then so are $A + BK$ and $C + DK$, respectively. Substituting $R = KQ$ in the matrix inequality and applying a congruence transformation gives us

$$Q^{-1}\left(Q(A+BK)^T + (A+BK)Q\right)Q^{-1} = (A+BK)^T Q^{-1} + Q^{-1}(A+BK) \prec 0.$$

Take the diagonal matrix $P \coloneqq Q^{-1} \succ 0$. Thus, by Theorem 3.9, we know that $\Sigma_{\mathrm{c}}^{\mathrm{cl}}$ is asymptotically stable. ■

For the discrete-time counterpart, we resort to statement (5) from Theorem 3.10 instead.

Theorem 5.2. *Given the system Σ_d, there exists a feedback gain K, which makes the closed-loop system Σ_{d}^{cl} positive and asymptotically stable if and only if there exist an $n \times n$ diagonal matrix $Q \succ 0$ and a matrix $R \in \mathbb{R}^{m \times n}$ such that $AQ + BR \geq 0$, $CQ + DR \geq 0$ and*

$$\begin{bmatrix} -Q & AQ+BR \\ (AQ+BR)^T & -Q \end{bmatrix} \prec 0.$$

Moreover, a feedback gain, which achieves the desired objective, is given by $K = RQ^{-1}$.

Proof. Similarly as in the proof of Theorem 5.1, we can show that the system is positive if and only if $AQ + BR \geq 0$ and $CQ + DR \geq 0$, so we only focus on the condition for asymptotic stability.

($\Rightarrow$): Assume there exists a K such that $\Sigma_{\mathrm{d}}^{\mathrm{cl}}$ is positive and asymptotically stable. By Theorem 3.10, there exists a diagonal $P \succ 0$ such that

$$(A+BK)^T P(A+BK) - P \prec 0.$$

By [12, Lemma A.2.1], for its Schur complement, it holds that

$$\begin{bmatrix} -P & P(A+BK) \\ (A+BK)^T P & -P \end{bmatrix} \prec 0.$$

We apply the following congruence transformation to obtain

$$\begin{aligned} \begin{bmatrix} P^{-1} & 0 \\ 0 & P^{-1} \end{bmatrix}^T \begin{bmatrix} -P & P(A+BK) \\ (A+BK)^T P & -P \end{bmatrix} \begin{bmatrix} P^{-1} & 0 \\ 0 & P^{-1} \end{bmatrix} &= \begin{bmatrix} -P^{-1} & (A+BK)P^{-1} \\ P^{-1}(A+BK)^T & -P^{-1} \end{bmatrix} \\ &= \begin{bmatrix} -Q & AQ+BR \\ (AQ+BR)^T & -Q \end{bmatrix} \prec 0, \end{aligned}$$

where we define $Q \coloneqq P^{-1} \succ 0$ and $R \coloneqq KQ$, i.e. $K = RQ^{-1}$.

($\Leftarrow$): Let $K = RQ^{-1}$. Then,

$$\begin{bmatrix} -Q & AQ+BR \\ (AQ+BR)^T & -Q \end{bmatrix} = \begin{bmatrix} -Q & (A+BK)Q \\ Q(A+BK)^T & -Q \end{bmatrix} \prec 0.$$

We apply the following congruence transformation:

$$\begin{bmatrix} Q^{-1} & 0 \\ 0 & Q^{-1} \end{bmatrix}^T \begin{bmatrix} -Q & (A+BK)Q \\ Q(A+BK)^T & -Q \end{bmatrix} \begin{bmatrix} Q^{-1} & 0 \\ 0 & Q^{-1} \end{bmatrix} = \begin{bmatrix} -Q^{-1} & Q^{-1}(A+BK) \\ (A+BK)^T Q^{-1} & -Q^{-1} \end{bmatrix} \prec 0.$$

After defining $P \coloneqq Q^{-1}$ and taking the Schur complement, we verify with Theorem 3.9 that $\Sigma_{\mathrm{d}}^{\mathrm{cl}}$ is asymptotically stable. ■

5.2 By Linear Programming

Statements (3) from Theorems 3.9 and 3.10 provide a method to convert SDP problems into LP ones. The following theorems theoretically yield the same controller design, though, Theorems 5.1 and 5.2 are more computationally expensive.

Theorem 5.3. *Given the system Σ_c, there exists a feedback gain K, which makes the closed-loop system Σ_c^{cl} positive and asymptotically stable if and only if there exist an $n \times n$ diagonal matrix $Q \succ 0$ and a matrix $R \in \mathbb{R}^{m \times n}$ such that $AQ + BR$ is Metzler, $CQ + DR \geq 0$ and*

$$(AQ + BR)\mathbb{1}_n < 0.$$

Moreover, a feedback gain, which achieves the desired objective, is given by $K = RQ^{-1}$.

Proof. Similarly as in the proof of Theorem 5.1, we can show that $A+BK$ is Metzler and $C+DK \geq 0$ if and only if $AQ+BR$ is Metzler and $CQ+DR \geq 0$, so we only focus on the condition for asymptotic stability.

(⇒): If $\Sigma_\mathrm{c}^\mathrm{cl}$ is asymptotically stable, then, by Theorem 3.9, there exists a $\mathbf{q} > 0$ such that $(A+BK)\mathbf{q} < 0$. By defining $Q \coloneqq \mathrm{diag}(\mathbf{q}) > 0$, we can rephrase the aforementioned as follows: there exists a diagonal matrix $Q \succ 0$ such that

$$(A + BK)\mathbf{q} = (A + BK)Q\mathbb{1}_n = (AQ + BKQ)\mathbb{1}_n < 0.$$

Taking $R \coloneqq KQ$ gives us the desired inequality.

(⇐): Let $K = RQ^{-1}$. Then,

$$(AQ + BR)\mathbb{1}_n = (A + BK)Q\mathbb{1}_n = (A + BK)\mathbf{q} < 0,$$

where $\mathbf{q} \coloneqq \mathrm{diag}(Q) > 0$. Therefore, by Theorem 3.9, we have that $A + BK$ is Hurwitz. ■

The discrete-time analogue of Theorem 5.3 can be easily formulated and proven.

Theorem 5.4. *Given the system Σ_d, there exists a feedback gain K, which makes the closed-loop system Σ_d^{cl} positive and asymptotically stable if and only if there exist an $n \times n$ diagonal matrix $Q \succ 0$ and a matrix $R \in \mathbb{R}^{m \times n}$ such that $AQ + BR \geq 0$, $CQ + DR \geq 0$ and*

$$(AQ - Q + BR)\mathbb{1}_n < 0.$$

Moreover, a feedback gain, which achieves the desired objective, is given by $K = RQ^{-1}$.

Proof. We can derive the proof by using the same reasoning as in the proof of Theorem 5.3, though we invoke Theorem 3.10 instead:

$$A + BK \text{ is Schur if and only if there exists a } \mathbf{q} > 0 \text{ such that } (A + BK)\mathbf{q} < \mathbf{q}.$$

■

5.3 A Numerical Example

Consider the following unstable discrete-time system, which is clearly not positive:

$$A = \begin{bmatrix} 0.2394 & 0.1327 & 0.2501 \\ -0.1752 & 1.4432 & 0.1018 \\ 0.9860 & -0.0663 & 0.3371 \end{bmatrix}, \; B = \begin{bmatrix} 0.0701 \\ 1.1232 \\ -1.1301 \end{bmatrix},$$

$$C = \begin{bmatrix} -0.1218 & 0.7426 & 0.6789 \end{bmatrix}, \; D = 1.0281.$$

After using YALMIP with MOSEK, we obtain from Theorem 5.2

$$
\begin{aligned}
K &= \begin{bmatrix} 0.2853 & -0.6698 & -0.0439 \end{bmatrix}, \\
\sigma(A+BK) &= \{-0.0702,\ 0.4543,\ 0.9529\}, \\
A+BK &= \begin{bmatrix} 0.2594 & 0.0857 & 0.2470 \\ 0.1453 & 0.6909 & 0.0524 \\ 0.6636 & 0.6907 & 0.3868 \end{bmatrix}, \\
C+DK &= \begin{bmatrix} 0.1715 & 0.0540 & 0.6337 \end{bmatrix}.
\end{aligned}
$$

Theorem 5.4 instead gives

$$
\begin{aligned}
K &= \begin{bmatrix} 0.2695 & -0.6441 & -0.0626 \end{bmatrix}, \\
\sigma(A+BK) &= \{-0.0665,\ 0.5130,\ 0.9394\}, \\
A+BK &= \begin{bmatrix} 0.2583 & 0.0876 & 0.2457 \\ 0.1274 & 0.7198 & 0.0315 \\ 0.6815 & 0.6616 & 0.4078 \end{bmatrix}, \\
C+DK &= \begin{bmatrix} 0.1552 & 0.0804 & 0.6146 \end{bmatrix}.
\end{aligned}
$$

Chapter 6

Data-Driven Analysis and Control

In this chapter, we extend the literature on positive systems by shifting to a data-based setting. More formally, we assume that the system of interest is positive, though, we have no a priori knowledge of the system matrices (A, B). Instead, we have access to finitely many input-state measurements. In Section 6.1, after introducing the framework for this chapter, we state necessary and sufficient conditions on the data, which would allow us to uniquely identify the system that generated the data. We demonstrate the result with two data sets in Subsection 6.1.1. Lastly, in Section 6.2, we set up future topics for research.

6.1 System Identification

The theoretical foundation of this chapter is the *informativity approach* [26, 27]. Consider a model class $\mathcal{M}$, which is a set of systems that contains our system of interest $\mathcal{S}$. Suppose that the "true" system is unknown, but also that we have access to a data set $\mathcal{D}$ generated by the system. Then, we can define the set of explaining systems $\Sigma_{\mathcal{D}} \subseteq \mathcal{M}$, i.e. systems that could have generated the data (see Figure 2).

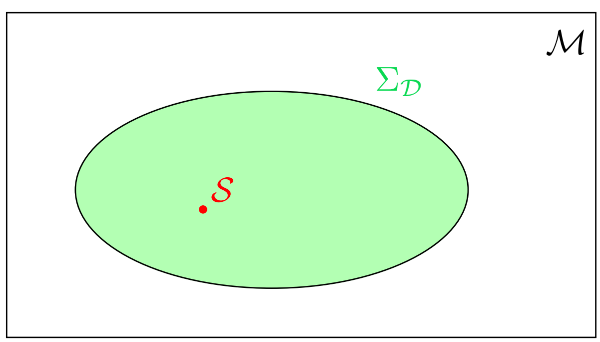

Figure 2: The system of interest $\mathcal{S}$ and the set of explaining systems $\Sigma_{\mathcal{D}}$.

In this section, we solve the problem of *system identification*. In other words, we seek to identify an a priori unknown system from data. To this end, we define informative data as follows.

Definition 6.1. The data $\mathcal{D}$ are called *informative for system identification* if $\Sigma_{\mathcal{D}} = \{\mathcal{S}\}$.

Naturally, we are interested in deriving necessary and sufficient conditions for informativity.

To formalise the problem formulation, consider the system

$$\mathbf{x}(t+1) = A_s\mathbf{x}(t) + B_s\mathbf{u}(t), \tag{6.1.1}$$

where A_s and B_s denote the true unknown system matrices. We collect input-state measurements in the matrices

$$U_- := \begin{bmatrix}\mathbf{u}(0) & \mathbf{u}(1) & \cdots & \mathbf{u}(T-1)\end{bmatrix},$$
$$X := \begin{bmatrix}\mathbf{x}(0) & \mathbf{x}(1) & \cdots & \mathbf{x}(T)\end{bmatrix}.$$

In addition, we define

$$X_- := \begin{bmatrix}\mathbf{x}(0) & \mathbf{x}(1) & \cdots & \mathbf{x}(T-1)\end{bmatrix},$$
$$X_+ := \begin{bmatrix}\mathbf{x}(1) & \mathbf{x}(2) & \cdots & \mathbf{x}(T)\end{bmatrix}.$$

Then, by noting (6.1.1), we see that the data matrices are related in the following way:

$$X_+ = AX_- + BU_- = \begin{bmatrix}A & B\end{bmatrix}\begin{bmatrix}X_- \\ U_-\end{bmatrix}, \tag{6.1.2}$$

where (A, B) is a system consistent with the data. Let all such *positive* systems be contained in

$$\Sigma_{\geq 0} := \{(A, B) \mid (6.1.2) \text{ holds and } A \geq 0,\ B \geq 0\}.$$

We, therefore, look for conditions on the data (U_-, X) such that $\Sigma_{\geq 0} = \{(A_s, B_s)\}$.

The system identification problem for *unconstrained* systems, i.e. systems in

$$\Sigma := \{(A, B) \mid (6.1.2) \text{ holds}\},$$

has already been studied in [26, Proposition 6].

Proposition 6.2. *The data (U_-, X), generated by $(A_s, B_s) \in \Sigma$, are informative for system identification if and only if*

$$\operatorname{rank}\begin{bmatrix}X_- \\ U_-\end{bmatrix} = n + m.$$

Intuitively, in the case of positive systems, we can weaken the assumption as it is not necessary that there is only one system consistent with the data. In fact, there can be many, but we require that exactly one of them is positive. We demonstrate this insight with the following example:

$$A_s = \begin{bmatrix}0 & 0 \\ 1 & 0\end{bmatrix},\ B_s = \begin{bmatrix}1 \\ 0\end{bmatrix},\ X = \begin{bmatrix}1 & 0 & 1 \\ 0 & 1 & 0\end{bmatrix},\ U_- = \begin{bmatrix}1 & 1\end{bmatrix}.$$

Then,

$$\begin{bmatrix}X_- \\ U_-\end{bmatrix} = \begin{bmatrix}1 & 0 \\ 0 & 1 \\ 1 & 1\end{bmatrix},\ X_+ = \begin{bmatrix}0 & 1 \\ 1 & 0\end{bmatrix}$$

and all system (A, B) that satisfy (6.1.2) are of the form

$$\left(\begin{bmatrix}-1+\alpha_1 & \alpha_1 \\ 1+\alpha_2 & \alpha_2\end{bmatrix}, \begin{bmatrix}1-\alpha_1 \\ -\alpha_2\end{bmatrix}\right)$$

for some $\alpha_1, \alpha_2 \in \mathbb{R}$. However, only one system is positive: $\alpha_1 = 1$ and $\alpha_2 = 0$.

Before stating and proving the positive systems version of [26, Proposition 6], we mention some notation that is to be used. Let (A, B) be *any* system in $\Sigma_{\geq 0}$. For $i = 1, \ldots, n$, define the sets

$$\mathcal{R}_i := \left\{j \mid \begin{bmatrix}A & B\end{bmatrix}_{ij} = 0\right\}. \tag{6.1.3}$$

Additionally, given a vector $\mathbf{v} \in \mathbb{R}^q$ and a set of indices $\mathcal{I} = \{i_1, \ldots, i_p\} \subseteq \{1, \ldots, q\}$, define the following vector:

$$\mathbf{v}_{\mathcal{I}} := \begin{bmatrix}v_{i_1} \\ v_{i_2} \\ \vdots \\ v_{i_p}\end{bmatrix} \in \mathbb{R}^p.$$

Theorem 6.3. *Let $(A, B) \in \Sigma_{\geq 0}$ and consider $\mathcal{R}_i$ in (6.1.3). The data (U_-, X), generated by $(A_s, B_s) \in \Sigma_{\geq 0}$, are informative for system identification if and only if, for every $i = 1, \ldots, n$,*

$$\begin{bmatrix} \boldsymbol{\xi}^T & \boldsymbol{\eta}^T \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} = 0 \quad \textit{and} \quad \begin{bmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{bmatrix}_{\mathcal{R}_i} \geq 0 \quad \textit{imply that} \quad \begin{bmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{bmatrix} = 0.$$

Proof. **(⇒):** Suppose there exists an i and a nonzero $\begin{bmatrix} \boldsymbol{\xi}^T & \boldsymbol{\eta}^T \end{bmatrix}$ such that

$$\begin{bmatrix} \boldsymbol{\xi}^T & \boldsymbol{\eta}^T \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} = 0 \quad \text{and} \quad \begin{bmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{bmatrix}_{\mathcal{R}_i} \geq 0.$$

Define a new system

$$\begin{aligned} \widehat{A} &\coloneqq A + \epsilon \mathbf{e}_i \boldsymbol{\xi}^T, \\ \widehat{B} &\coloneqq B + \epsilon \mathbf{e}_i \boldsymbol{\eta}^T, \end{aligned}$$

where $\epsilon > 0$ is a scalar. Evidently, $(\widehat{A}, \widehat{B})$ satisfy (6.1.2). Next, we show that $\widehat{A} \geq 0$ and $\widehat{B} \geq 0$ by studying the ith row of $\begin{bmatrix} \widehat{A} & \widehat{B} \end{bmatrix}$. Firstly, if $j \in \mathcal{R}_i$, then

$$\begin{bmatrix} \widehat{A} & \widehat{B} \end{bmatrix}_{ij} = \epsilon \begin{bmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{bmatrix}_j \geq 0.$$

Secondly, consider all j, for which $\begin{bmatrix} A & B \end{bmatrix}_{ij} > 0$. Note that we can choose a sufficiently small ϵ such that $\begin{bmatrix} \widehat{A} & \widehat{B} \end{bmatrix}_{ij} \geq 0$. Thus, $(\widehat{A}, \widehat{B}) \in \Sigma_{\geq 0}$. However, $(\widehat{A}, \widehat{B}) \neq (A, B)$, so there are at least two explaining systems, which leads to a contradiction.

(⇐): Consider the two systems $(A, B), (\tilde{A}, \tilde{B}) \in \Sigma_{\geq 0}$. Because both systems satisfy (6.1.2), we have that

$$\begin{bmatrix} \tilde{A} - A & \tilde{B} - B \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} = 0.$$

Denote the ith row of the matrix on the left by

$$\mathbf{c}_i^T \coloneqq \begin{bmatrix} \tilde{A} - A & \tilde{B} - B \end{bmatrix}_{i\bullet},$$

where $\mathbf{c}_1, \ldots, \mathbf{c}_n$ are vectors. We know that, for all $i = 1, \ldots, n$ and $j \in \mathcal{R}_i$,

$$\begin{bmatrix} \tilde{A} - A & \tilde{B} - B \end{bmatrix}_{ij} = \begin{bmatrix} \tilde{A} & \tilde{B} \end{bmatrix}_{ij} \geq 0,$$

i.e. $\left(\mathbf{c}_i^T\right)_{\mathcal{R}_i} \geq 0$. By assumption, this implies that $\mathbf{c}_i^T = 0$ for all i. Hence, $\begin{bmatrix} \tilde{A} - A & \tilde{B} - B \end{bmatrix} = 0$ and, consequently, $(A, B) = (\tilde{A}, \tilde{B}) = (A_s, B_s)$. ∎

Note that data are neither required to be sufficiently rich nor to be nonnegative for them to be informative (see Subsection 6.1.1). Based on Theorem 6.3, the procedure of verifying informativity for system identification would go as follows:

1. Generate a positive system $(\bar{A}, \bar{B})$ row-wise from the data using LP;
2. Define the sets $\mathcal{R}_i$ $(i = 1, \ldots, n)$ with respect to $(\bar{A}, \bar{B})$;
3. Check whether there are numbers with opposite signs among the elements, with indices in $\mathcal{R}_i$, of nonzero vectors in the left kernel of $\begin{bmatrix} X_-^T & U_-^T \end{bmatrix}^T$.

In the end, if the data are informative for system identification, then $(\bar{A}, \bar{B}) = (A_s, B_s)$. We conclude this section with a straightforward consequence of Theorem 6.3.

Corollary 6.4. *Assume that $A_s > 0$ and $B_s > 0$. The data (U_-, X), generated by (A_s, B_s), are informative for system identification if and only if*

$$\operatorname{rank}\begin{bmatrix} X_- \\ U_- \end{bmatrix} = n + m.$$

Proof. By the definition of the sets $\mathcal{R}_i$, we have that $\mathcal{R}_i = \varnothing$ for all $i = 1, \ldots, n$, so the condition

$$\begin{bmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{bmatrix}_{\mathcal{R}_i} \geq 0$$

is trivially satisfied. Therefore, by Theorem 6.3, only the zero vector is in the left kernel of $\begin{bmatrix} X_-^T & U_-^T \end{bmatrix}^T$ or, equivalently, $\begin{bmatrix} X_-^T & U_-^T \end{bmatrix}^T$ is full row rank. ■

6.1.1 Numerical Examples

Consider the two data sets

$$\mathcal{D}_1 = \left\{ U_- = \begin{bmatrix} -29.4678 \\ -18.5209 \\ 5.2541 \end{bmatrix}^T, \; X = \begin{bmatrix} -1.7535 & -59.5498 & -4.4944 & -55.9726 \\ 1.9781 & 8.5859 & -118.5480 & -843.3192 \\ -0.1535 & 8.1369 & -16.6202 & -597.2345 \end{bmatrix} \right\},$$

$$\mathcal{D}_1 = \left\{ U_- = \begin{bmatrix} 6.7842 & -5.0283 \\ 5.2211 & -1.2966 \\ -3.0960 & -1.7551 \\ 10.6810 & -3.0800 \end{bmatrix}^T, \; X = \begin{bmatrix} -1 & 4.4 & 1.2 & -210.4 & -1783.3 \\ -1.5 & 9.8 & -62.4 & 71.8 & -1364 \\ -0.6 & 3.3 & 96.9 & -40.5 & -596.2 \\ -0.3 & -22.5 & -48 & -177.9 & -2537.4 \end{bmatrix} \right\}.$$

We remark that $\begin{bmatrix} X_-^T & U_-^T \end{bmatrix}^T$ is not full row rank for neither data set. By using MATLAB's Optimisation Toolbox, we generate from $\mathcal{D}_1$ the positive system

$$\bar{A} = \begin{bmatrix} 0 & 0 & 4 \\ 3 & 7 & 0 \\ 1 & 5 & 0 \end{bmatrix}, \; \bar{B} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}.$$

Then, we have

$$\mathcal{R}_1 = \{1,\ 2\},$$
$$\mathcal{R}_2 = \{3,\ 4\},$$
$$\mathcal{R}_3 = \{3,\ 4\}.$$

The command *null*() returns the following orthonormal basis for the left kernel of $\begin{bmatrix} X_-^T & U_-^T \end{bmatrix}^T$:

$$\left\{ \begin{bmatrix} -0.1204 \\ 0.1432 \\ -0.9821 \\ 0.0219 \end{bmatrix} \right\}.$$

Because the first two and the last two elements have opposite signs, we see that the data in $\mathcal{D}_1$ are informative for system identification and $(\bar{A}, \bar{B})$ is the true system.

Similarly, from $\mathcal{D}_2$, we obtain

$$\bar{A} = \begin{bmatrix} 7.0114 & 1.5009 & 0 & 2.4670 \\ 2.4950 & 0 & 2.8617 & 4.1996 \\ 4.2369 & 5.5017 & 3.1558 & 0 \\ 8.8776 & 0.0682 & 0 & 3.7494 \end{bmatrix}, \; \bar{B} = \begin{bmatrix} 2.1501 & 0 \\ 2.2467 & 0 \\ 2.6173 & 0 \\ 0 & 2.4153 \end{bmatrix}, \; \begin{aligned} \mathcal{R}_1 &= \{3,\ 6\}, \\ \mathcal{R}_2 &= \{2,\ 6\}, \\ \mathcal{R}_3 &= \{4,\ 6\}, \\ \mathcal{R}_3 &= \{3,\ 5\}. \end{aligned}$$

An orthonormal basis for the left kernel of $\begin{bmatrix} X_-^T & U_-^T \end{bmatrix}^T$ is

$$\left\{ \begin{bmatrix} 0.1405 \\ -0.5884 \\ -0.5056 \\ -0.2846 \\ 0.2113 \\ 0.5027 \end{bmatrix}, \begin{bmatrix} -0.1509 \\ 0.2967 \\ 0.3421 \\ 0.2424 \\ 0.5533 \\ 0.6381 \end{bmatrix} \right\},$$

but the second vector does not satisfy the conditions in Theorem 6.3. Thus, the data in $\mathcal{D}_2$ are not informative for system identification. The system that we used to generate $\mathcal{D}_2$ is

$$A_s = \begin{bmatrix} 6 & 3 & 2 & 4 \\ 2 & 1 & 4 & 5 \\ 3 & 8 & 6 & 2 \\ 7 & 5 & 5 & 7 \end{bmatrix}, \; B_s = \begin{bmatrix} 7 & 6 \\ 4 & 2 \\ 7 & 5 \\ 4 & 6 \end{bmatrix},$$

so we can also conclude from Corollary 6.4 that the data are not informative.

6.2 On the Stability and Stabilisation Problems

Suppose we are interested in a system property $\mathcal{P}$ or a control objective $\mathcal{O}$. In data-driven analysis problems, we want to verify whether $\mathcal{S}$ has property $\mathcal{P}$. But because $\mathcal{S}$ is, essentially, indistinguishable from the rest of the systems in $\Sigma_{\mathcal{D}}$, we need to ensure that all systems in $\Sigma_{\mathcal{D}}$ have property $\mathcal{P}$. In data-driven control problems, we seek a data-based design for a controller $\mathcal{K}$, which, when interconnected with $\mathcal{S}$, makes the closed-loop system satisfy objective $\mathcal{O}$. Similarly, to guarantee this, we want all systems in $\Sigma_{\mathcal{D}}$ to satisfy objective $\mathcal{O}$ when interconnected with $\mathcal{K}$. This leads to the following definitions.

Definition 6.5. The data $\mathcal{D}$ are called *informative for property* $\mathcal{P}$ if all systems in $\Sigma_{\mathcal{D}}$ have property $\mathcal{P}$.

Definition 6.6. The data $\mathcal{D}$ are called *informative for objective* $\mathcal{O}$ if there exists a controller $\mathcal{K}$ such that all systems in $\Sigma_{\mathcal{D}}$ satisfy objective $\mathcal{O}$ when interconnected with $\mathcal{K}$.

More specifically, we can refer to $\mathcal{P}$ and $\mathcal{O}$ as stability and positive stabilisation, respectively. Firstly, for the data-driven stability problem, we are interested in verifying the stability of the autonomous positive system

$$\mathbf{x}(t+1) = A_s\mathbf{x}(t). \tag{6.2}$$

We collect state data in the matrix X as defined in the previous section. This time, the set of systems that are consistent with the data is

$$\Sigma_{\geq 0}^{\text{stab}} := \{A \colon X_+ = AX_- \text{ and } A \geq 0\}.$$

Then, we define informativity for stability as follows.

Definition 6.7. The data X, generated by (6.2), are called *informative for stability* if all matrices in $\Sigma_{\geq 0}^{\text{stab}}$ are Schur.

Secondly, for the data-driven positive stabilisation problem, we want to find a controller $\mathbf{u}(t) = K\mathbf{x}(t)$, which renders (6.1.1) positive and asymptotically stable. Because we do not require the system to be a priori positive, the set of explaining systems is Σ as defined in the previous section. The following is the definition of informativity for positive stabilisation by state feedback.

Definition 6.8. The data (U_-, X), generated by (6.1.1), are called *informative for positive stabilisation by state feedback* if there exists a feedback gain K such that $A + BK$ is nonnegative Schur for all $(A, B) \in \Sigma$.

Deriving necessary and sufficient conditions for informativity for stability and positive stabilisation as well as providing a data-based design for K remain open problems. To conclude, we note that in this chapter we have only considered *noiseless* data, so an additional research topic is to devise conditions for informativity, while dealing with *noisy* measurements.

Chapter 7

Conclusion

In this thesis, we have created an overview of the fundamental results in positive systems theory. In addition, we have expanded the literature on positive systems by moving to a data-driven setting and solving the system identification problem. Alongside the statements, we have also added accessible proofs. More explicitly, we have done the following.

- We have characterised both continuous- and discrete-time positive systems in Theorem 3.4.
- We have shown 5 methods for verifying asymptotic stability in Theorems 3.6 and 3.7. As a consequence, these theorems have allowed us to reformulate SDP problems for verifying stability of positive systems into LP ones, leading to a reduction in computational complexity. Furthermore, the theorems have shown that working with positive systems in the context of large-scale systems remains practically feasible.
- We have stated and proven the Kalman-Yakubovich-Popov lemma for positive systems (Theorems 4.1 and 4.2), which links frequency domain conditions and the solvability of an LMI. We have also seen that it is enough to verify the frequency domain condition for the zero frequency only.
- We have shown necessary and sufficient conditions for the existence of a controller, which makes a (not necessarily positive) system stable and positive. Additionally, we have provided a controller design, which can be obtained by using SDP (Theorems 5.1 and 5.2) or LP (Theorems 5.3 and 5.4).
- We have derived necessary and sufficient conditions for data to be informative for system identification in Theorem 6.3. Moreover, we have provided specific steps for verifying informativity, which have been demonstrated with numerical examples.

Interesting topics for future research are the following.

- Based on the setup in Section 6.2 of Chapter 6, we would like to solve the data-driven stability and positive stabilisation problems by using the concept of informative data.
- As we have only considered noiseless data, a natural extension would include the use of noisy data.
- Positive systems are such that if a trajectory enters the positive orthant of $\mathbb{R}^n$, then it stays there, provided that we apply nonnegative inputs. In other words, the positive orthant can be thought of as an invariant set. We would like to extend the theory of positive systems by considering certain classes of convex cones, instead of the positive orthant.

Bibliography

[1] D. G. Luenberger, *Introduction to Dynamic Systems: Theory, Models, and Applications.* New York, NY: John Wiley & Sons, 1979.

[2] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, ser. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 1994.

[3] L. Farina and S. Rinaldi, *Positive Linear Systems: Theory and Applications*, ser. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2000.

[4] T. Kaczorek, *Positive 1D and 2D Systems*, ser. Communications and Control Engineering. London: Springer, 2002.

[5] A. Rantzer and M. E. Valcher, "A tutorial on positive systems and large scale control," in *2018 IEEE Conference on Decision and Control (CDC)*, Miami, FL, USA, 2018, pp. 3686–3697.

[6] M. Humi, *Introduction to Mathematical Modeling.* Boca Raton, FL: CRC Press, Taylor & Francis Group, 2017.

[7] D. S. Bernstein, *Scalar, Vector, and Matrix Mathematics: Theory, Facts, and Formulas - Revised and Expanded Edition.* Princeton University Press, 2018.

[8] O. Perron, "Zur theorie der matrices," *Mathematische Annalen*, vol. 64, no. 2, pp. 248–263, 1907.

[9] F. G. Frobenius, *Über Matrizen aus nicht negativen Elementen.* Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, 1912, pp. 456–477.

[10] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, 1st ed., ser. Other Titles in Applied Mathematics. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2000, vol. 71.

[11] P. J. Antsaklis and A. N. Michel, *A Linear Systems Primer.* Boston, MA: Birkhäuser, 2007.

[12] R. E. Skelton, T. Iwasaki, and K. Grigoriadis, *A Unified Algebraic Approach to Linear Control Design.* London: Routledge, 1998.

[13] J. Löfberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *2004 IEEE International Conference on Robotics and Automation*, Taipei, Taiwan, 2004, pp. 284–289.

[14] MOSEK ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 9.0.*, 2019. [Online]. Available: http://docs.mosek.com/9.0/toolbox/index.html

[15] V. A. Yakubovich, "Solution of some matrix inequalities encountered in the automatic control theory," *Dokl. Akad. Nauk SSSR*, vol. 143, no. 6, pp. 1304–1307, 1962.

[16] R. E. Kalman, "Lyapunov functions for the problem of Lur'e in automatic control," in *Proceedings of the National Academy of Sciences of the United States of America*, vol. 49, no. 2, 1963, pp. 201–205.

[17] V. M. Popov, "Hyperstability and optimality of automatic systems with several control functions," *Rev. Roumaine Sci. Tech., Ser. Electrotech. Energ.*, vol. 9, no. 4, pp. 629–890, 1964.

[18] S. V. Gusev and A. L. Likhtarnikov, “Kalman-Popov-Yakubovich lemma and the S-procedure: A historical essay,” *Automation and Remote Control*, vol. 67, no. 11, pp. 1768–1810, 2006.

[19] A. Rantzer, “On the Kalman-Yakubovich-Popov lemma for positive systems,” in *Proceedings of 51st IEEE Conference on Decision and Control*, Maui, Hawaii, USA, 2012, pp. 7482–7484.

[20] ——, “An extended Kalman-Yakubovich-Popov lemma for positive systems,” *IFAC-PapersOnLine*, vol. 48, no. 11, pp. 242–245, 2015.

[21] ——, “On the Kalman-Yakubovich-Popov lemma for positive systems,” *IEEE Transactions on Automatic Control*, vol. 61, no. 5, pp. 1346–1349, 2016.

[22] V. Balakrishnan and L. Vandenberghe, “Semidefinite programming duality and linear time-invariant systems,” *IEEE Transactions on Automatic Control*, vol. 48, no. 1, pp. 30–41, 2003.

[23] T. Tanaka and C. Langbort, “KYP lemma for internally positive systems and a tractable class of distributed H-infinity control problems,” in *Proceedings of the 2010 American Control Conference*, Baltimore, MD, USA, 2010, pp. 6238–6243.

[24] H. Gao, J. Lam, C. Wang, and S. Xu, “Control for stability and positivity: Equivalent conditions and computation,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 52, no. 9, pp. 540–544, 2005.

[25] M. A. Rami and F. Tadeo, “Controller synthesis for positive linear systems with bounded controls,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 54, no. 2, pp. 151–155, 2007.

[26] H. J. van Waarde, J. Eising, H. L. Trentelman, and M. K. Camlibel, “Data informativity: A new perspective on data-driven analysis and control,” *IEEE Transactions on Automatic Control*, vol. 65, no. 11, pp. 4753–4768, 2020.

[27] H. J. van Waarde, J. Eising, M. K. Camlibel, and H. L. Trentelman, “The informativity approach: To data-driven analysis and control,” *IEEE Control Systems*, vol. 43, no. 6, pp. 32–66, 2023.