29th Internet Seminar

Eventual Positivity

Sahiba Arora, Jochen Glück, Jonathan Mui

Lecture Notes
Winter term 2025/26

Contents

Introduction

Prerequisites

This course was designed for postgraduate students (Masters and PhD) and advanced Bachelor students with the following prerequisite knowledge:

- Calculus/analysis in one and several variables;
- Linear algebra (in particular eigenvalues and the Jordan normal form of matrices);
- An introduction to real analysis, measure and integration theory (in particular, familiarity with L^p spaces);
- An introduction to functional analysis (Banach spaces, Hilbert spaces, bounded linear operators); and
- An introduction to complex analysis (holomorphic functions, complex path integrals and Cauchy's integral formula, Laurent series).

Acknowledgements

We thank Alexander Wierzba for sending us the font used in these lecture notes.

Nomenclature

The following table gives an overview of important symbols used in the lectures. This table will be updated every week as we introduce new notation.

Elementary notation

$\mathbb{N}$	strictly positive integers, $\mathbb{N} = \{1,2,3,\dots\}$
$\mathbb{N}_0$	non-negative integers, $\mathbb{N}_0 = \{0,1,2,\dots\}$
$\mathbb{R}_+$	alternative notation for the interval $[0,\infty)$

Function spaces

$\mathbb{1}$	The vector in $\mathbb{R}^n$ whose entries are all 1, or the constant function with value 1 on a set that is clear from the context
id	The identity matrix in $\mathbb{C}^{n\times n}$ or the identity operator on a normed space that is clear from the context

Spectral theory

$\sigma(A)$	spectrum of linear operator A
$\rho(A)$	resolvent set of a linear operator A, i.e. $\rho(A) := \mathbb{C} \setminus \sigma(A)$
$\mathcal{R}(\lambda, A)$	resolvent of A at a point $\lambda \in \rho(A)$, i.e. $\mathcal{R}(\lambda, A) := (\lambda - A)^{-1}$
$\mathrm{r}(A)$	spectral radius of a bounded linear operator A, defined by the formula $\mathrm{r}(A) := \max\{\lvert\lambda\rvert : \lambda \in \sigma(A)\} \in [0,\infty)$
$\mathrm{s}(A)$	spectral bound of a linear operator A, defined by the formula $\mathrm{s}(A) := \sup\{\operatorname{Re}\lambda : \lambda \in \sigma(A)\} \in [-\infty,\infty]$

Ordered structures on vector spaces

$x \preceq y$	$cx \geq y$ for some number $c > 0$ (equivalently, $x \leq cy$ for some number $c > 0$)
$y \succeq x$	alternative notation for $x \preceq y$

Chapter 1

Positive matrices and matrix semigroups

The topic of the ISEM 29 is the interplay between dynamical systems (more specifically: differential equations), sign preservation, and operator theory. The material in the first two chapters develops the essence of the theory in finite dimensions. In Chapter 1, we study positive matrices and matrix exponential functions, and show how the positivity affects their eigenvalues and eigenvectors. The titular subject, *eventual positivity*, makes an appearance in Chapter 2.

1.1 Positive matrices and the standard order on $\mathbb{R}^n$

As a foundation for everything that follows, we endow the spaces $\mathbb{R}^n$ and $\mathbb{R}^{m\times n}$ with the following partial order.

Definition 1.1.1 (The order and the cone on $\mathbb{R}^n$ and $\mathbb{R}^{m\times n}$)**.**

(a) For $x, y \in \mathbb{R}^n$ we write $x \le y$ if this inequality holds componentwise, i.e. if $x_k \le y_k$ for every index k. As usual we use the notation $y \ge x$ synomymously with $x \le y$.

Vectors $x \in \mathbb{R}^n$ that satisfy $x \ge 0$ [1] are called the **positive** elements of $\mathbb{R}^n$, and the set $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n : x \ge 0\}$ of all positive vectors is called the **positive cone** of $\mathbb{R}^n$.

(b) We use the same conventions for matrices: for $A, B \in \mathbb{R}^{m\times n}$ we write $A \le B$ (or $B \ge A$) if $A_{jk} \le B_{jk}$ for all indices j, k.

A matrix $A \in \mathbb{R}^{m\times n}$ is called **positive** if $A \ge 0$, and the set $\mathbb{R}^{m\times n}_+ := \{A \in \mathbb{R}^{m\times n} : A \ge 0\}$ of positive matrices is called the **positive cone** in $\mathbb{R}^{m\times n}$.

Note that Definition 1.1.1(a) can be considered a special case of part (b) if we identify $\mathbb{R}^n$ with $\mathbb{R}^{n\times 1}$. The relation $\le$ is a partial order on $\mathbb{R}^n$ and is compatible with its vector

[1] We always write 0 for the zero vector when the corresponding space is clear from context.

space structure in the following sense: if $x \le y$ for $x, y \in \mathbb{R}^n$, then

$$\alpha x \le \alpha y \qquad \text{and} \qquad x + z \le y + z$$

for all numbers $\alpha \in [0, \infty)$ and all vectors $z \in \mathbb{R}^n$. Analogous statements hold for the partial order $\le$ on $\mathbb{R}^{m\times n}$.

Remark 1.1.2 (Terminology: positive vectors)**.** At first glance, it might be surprising that our definition of 'positivity' is inconsistent with its common meaning for real numbers: in English, a number $\alpha \in \mathbb{R}$ is usually called positive if $\alpha > 0$. Yet, the vector $0 \in \mathbb{R}^n$ is positive in the sense of Definition 1.1.1. For $n = 1$ this means, in particular, that the real number 0 is positive in the sense of Definition 1.1.1.

Nevertheless, our usage of 'positive' is standard in the theory of Banach lattices, which we use frequently from Chapter 4 on. For readers who take pleasure in terminological digressions, a more thorough discussion is provided in the notes at the end of this chapter.

From an operator-theoretic perspective, it is desirable to describe positivity of matrices in terms of how they act as linear maps. We do this in the next proposition.

Proposition 1.1.3. *For a matrix $A \in \mathbb{R}^{m\times n}$, the following are equivalent:*

(i) *A is positive, i.e. $A \in \mathbb{R}^{m\times n}_+$.*

(ii) *$A(\mathbb{R}^n_+) \subseteq \mathbb{R}^m_+$.*

(iii) *A acts monotonically, i.e. if $x, y \in \mathbb{R}^n$ satisfy $x \le y$, then $Ax \le Ay$.*

Proof. *"(i) ⇒ (ii)":* This is clear from the definition of the matrix-vector product.

"(ii) ⇒ (iii)": Assume that (ii) holds and let $x, y \in \mathbb{R}^n$ satisfy $x \le y$. Then $y - x \in \mathbb{R}^n_+$ and hence $Ay - Ax = A(y - x) \in \mathbb{R}^m_+$, which implies that $Ax \le Ay$.

"(iii) ⇒ (i)": Assume that (iii) holds. For $j \in \{1, \dots, n\}$ and the canonical unit vector $e_j \in \mathbb{R}^n$ one has $0 \le e_j$ and thus $0 = A0 \le Ae_j$. Since Ae_j is the j-th column of A and j was arbitrary, we conclude that all entries of A are ≥ 0. □

Since we defined the order relation $\le$ by comparing vectors (and matrices) entrywise, it is natural to generalise the modulus from scalars to vectors in the same way:

Definition 1.1.4 (The modulus of vectors and matrices)**.** For every vector $x \in \mathbb{C}^n$ and every matrix $A \in \mathbb{C}^{m\times n}$ we define the matrix $|A| \in \mathbb{R}^{m\times n}_+$ and the vector $|x| \in \mathbb{R}^n$ by taking the entrywise modulus of x and A, i.e.

$$|x|_j := |x_j| \qquad \text{and} \qquad |A|_{jk} := |A_{jk}|$$

for all indices j and k.

The modulus has a submultiplicative property, which is very useful to prove estimates for positive matrices.

Proposition 1.1.5 (Submultiplicativity of the modulus)**.** *Let* $A \in \mathbb{C}^{m\times n}$ *and* $x \in \mathbb{C}^n$.

(a) *One has* $|Ax| \le |A|\,|x|$.

(b) *In particular, if* $A \in \mathbb{R}_+^{m\times n}$, *then* $|Ax| \le A|x|$.

Proof. (a) One can check the inequality entrywise: for every $j \in \{1,\dots,n\}$ one has

$$|Ax|_j = \left|(Ax)_j\right| = \left|\sum_{k=1}^n A_{jk}x_k\right| \le \sum_{k=1}^n \left|A_{jk}\right| |x_k| = \big(|A|\,|x|\big)_j .$$

(b) For positive A, one has $|A| = A$, so the claim follows from part (a). □

Remark 1.1.6 (Norms on $\mathbb{C}^n$)**.** In the following we will often work with norms on $\mathbb{C}^n$. While they are all equivalent, we will assume throughout that $\mathbb{C}^n$ is endowed with a norm that satisfies $\||x|\| = \|x\|$ for all $x \in \mathbb{C}^n$ as well as $\|x\| \le \|y\|$ for all $x, y \in \mathbb{R}^n$ with $0 \le x \le y$ – this is sometimes more convenient in estimates. For instance, the p-norm has this property for every $p \in [1,\infty]$.

1.2 The spectrum of positive matrices

An intriguing feature of positive matrices is that their eigenvalues and eigenvectors enjoy a variety of remarkable properties. This is the content of the classical **Perron-Frobenius theorem**, which we study in this section. This theorem is only a first instance of one of the most important themes of the course: the interaction between positivity and the spectrum of linear operators. We take this opportunity to introduce some fundamental concepts and tools in spectral theory.

Definition 1.2.1 (Spectrum and spectral radius)**.** Let $A \in \mathbb{C}^{n\times n}$. The set $\sigma(A) \subseteq \mathbb{C}$ that consists of all eigenvalues of A is called the **spectrum** of A, and the number

$$\mathrm{r}(A) := \max\big\{|\lambda| : \lambda \in \sigma(A)\big\} \in [0,\infty)$$

is called the **spectral radius** of A.

The spectral radius determines whether the powers of a matrix converge to 0 as the exponent tends to ∞. More precisely, one has the following equivalence.

Proposition 1.2.2 (Convergence to 0 of matrix powers)**.** *For every matrix* $A \in \mathbb{C}^{n\times n}$, *the following are equivalent:*

(i) $\mathrm{r}(A) < 1$.

(ii) $A^k \to 0$ *as* $k \to \infty$.

(iii) *There exist numbers* $\eta \in [0,1)$ *and* $c \ge 0$ *such that* $\left\|A^k\right\| \le c\eta^k$ *for each* $k \in \mathbb{N}_0$.

Proof. *"(i) ⇒ (iii)":* The implication is clear if $\mathrm{r}(A) = 0$, hence we assume $\mathrm{r}(A) > 0$. One can then show, using the Jordan normal form of A, that $\|A^k\| \le \tilde{c}\,\mathrm{r}(A)^k(1+k^{n-1})$ for a number $\tilde{c} \ge 0$ and all $k \in \mathbb{N}_0$; see Exercise 1.4(c). So the claim follows by taking any $\eta \in (\mathrm{r}(A), 1)$ and using that $\frac{\mathrm{r}(A)^k}{\eta^k}$ decays exponentially.

"(iii) ⇒ (ii)": This implication is obvious.

"(ii) ⇒ (i)": Let λ be an eigenvalue of A with $|\lambda| = \mathrm{r}(A)$ associated to an eigenvector z of norm one. One has $|\lambda|^k = |\lambda|^k \|z\| = \|A^k z\| \to 0$ as $k \to \infty$. Thus, $\mathrm{r}(A) = |\lambda| < 1$. □

To formulate the statement of some parts of the Perron-Frobenius theorem we need the following stronger notion of positivity.

Definition 1.2.3 (Strong positivity in finite dimensions)**.** A vector $x \in \mathbb{R}^n$ is called **strongly positive** if $x_k > 0$ for all $k \in \{1, \dots, n\}$. Similarly, a matrix $A \in \mathbb{R}^{m\times n}$ is called **strongly positive** if $A_{jk} > 0$ for all indices j, k.

Observe that the strongly positive vectors in $\mathbb{R}^n$ are precisely the points in the interior of the positive cone $\mathbb{R}^n_+$. Similarly as in Proposition 1.1.3, strong positivity of matrices can also be interpreted in terms of their actions as linear mappings: a matrix $A \in \mathbb{R}^{m\times n}$ is strongly positive if and only if it maps every $0 \ne x \in \mathbb{R}^n_+$ to a strongly positive vector.

It is convenient to have a notation for strong positivity. The following has the advantage that it can easily be generalised to the infinite-dimensional setting in later chapters.

Notation 1.2.4 (Inequality up to a factor)**.**

(a) For two vectors $x, y \in \mathbb{R}^n$ we write

$$x \preceq y \qquad \text{or alternatively} \qquad y \succeq x$$

if there exists a number $c > 0$ such that $cx \le y$ (equivalently, if there exists a number $c > 0$ such that $x \le cy$).

(b) We let $\mathbb{1} \in \mathbb{R}^n$ denote the vector with every entry equal to 1. Hence, a vector $x \in \mathbb{R}^n$ is strongly positive if and only if $x \succeq \mathbb{1}$.

The main result of this section is the following classical theorem about the eigenvalues and eigenvectors of positive matrices.

Theorem 1.2.5 (Perron–Frobenius)**.** *Let* $0 \le A \in \mathbb{R}^{n\times n}$.

(a) *The spectral radius* $\mathrm{r}(A)$ *is an eigenvalue of* A *with an eigenvector* $x \ge 0$.

(b) *If all diagonal entries of* A *are non-zero, then* $\mathrm{r}(A) > 0$*, and* $\mathrm{r}(A)$ *is a* ***radially strictly dominant*** *eigenvalue in the sense that* $|\lambda| < \mathrm{r}(A)$ *for all other eigenvalues* λ *of* A.

(c) *If* A *is even strongly positive, then* $\mathrm{r}(A) > 0$*, the eigenvalue* $\mathrm{r}(A)$ *of* A *is algebraically simple, and its eigenspace is spanned by a strongly positive vector.*

The Perron–Frobenius theorem is quite useful to study the behaviour of the powers A^k of a positive matrix A as $k \to \infty$. We explore a concrete application to Markov chains in Exercise 1.3.

Various proofs of the theorem and variations thereof are known; see e.g. the survey article [Mac00] for some nice bedtime reading. The proof we present has a strong functional analytic flavour and already anticipates several ideas and arguments that will occur again in the infinite-dimensional case – strongly relying on properties of the resolvent. We define and study this object now and finally use it to prove Theorem 1.2.5.

Definition 1.2.6 (The resolvent of a matrix)**.** Let $A \in \mathbb{C}^{n\times n}$. The complement of its spectrum, i.e. $\rho(A) := \mathbb{C} \setminus \sigma(A)$, is called the **resolvent set** of A. The mapping

$$\mathcal{R}(\cdot, A)\colon \rho(A) \to \mathbb{C}^{n\times n}, \qquad \lambda \mapsto \mathcal{R}(\lambda, A) := (\lambda - A)^{-1}$$

is called the **resolvent** of A.

In the preceding definition, we used the notation $\lambda - A$, which is shorthand for $\lambda\,\mathrm{id} - A$; where $\mathrm{id} \in \mathbb{C}^{n\times n}$, denotes the identity matrix of the same dimension as A.

To state the following proposition we need the concept of a vector-valued analytic functions. In finite dimensions this is easy: a mapping from an open subset of $\mathbb{C}$ to $\mathbb{C}^n$ or to $\mathbb{C}^{n\times n}$ is called **analytic** or **holomorphic** if every component of the mapping is analytic.

Proposition 1.2.7 (Properties of resolvents)**.** *Let $A \in \mathbb{C}^{n\times n}$.*

(a) *The resolvent $\mathcal{R}(\cdot, A)\colon \rho(A) \to \mathbb{C}^{n\times n}$ is analytic.*

(b) *For $\lambda \in \mathbb{C}$ with $|\lambda| > \mathrm{r}(A)$, the resolvent can be represented as the* ***Neumann series***

$$\mathcal{R}(\lambda, A) = \sum_{k=0}^{\infty} \frac{A^k}{\lambda^{k+1}},$$

which converges absolutely in $\mathbb{C}^{n\times n}$ (with respect to any norm).

Proof. (a) It follows from Cramer's rule for the inverse of a matrix that, for all indices j, k, the matrix entry $\mathcal{R}(\cdot, A)_{jk}\colon \rho(A) \to \mathbb{C}$ is a rational function and thus analytic.

(b) As $\mathrm{r}(A/\lambda) < 1$, so by Proposition 1.2.2, there exist numbers $\eta \in [0, 1)$ and $c \geq 0$ such that $\left\| A^k/\lambda^k \right\| \leq c\eta^k$ for every $k \in \mathbb{N}_0$. Thus, $\sum_{k=0}^{\infty} \left\| \frac{A^k}{\lambda^{k+1}} \right\| < \infty$, and hence the series converges absolutely in $\mathbb{C}^{n\times n}$. To show the resolvent formula, we compute

$$(\lambda - A) \sum_{k=0}^{\infty} \frac{A^k}{\lambda^{k+1}} = \lim_{K\to\infty} \sum_{k=0}^{K} \left(\frac{A^k}{\lambda^k} - \frac{A^{k+1}}{\lambda^{k+1}} \right) = \lim_{K\to\infty} \left(\mathrm{id} - \frac{A^{K+1}}{\lambda^{K+1}} \right) = \mathrm{id}.$$

Here we used that $\left\| A^{k+1}/\lambda^{k+1} \right\| \to 0$ as $k \to \infty$ according to Proposition 1.2.2 since $\mathrm{r}(A/\lambda) < 1$. After multiplying by $\mathcal{R}(\lambda, A)$, we obtain the claimed formula. □

Finally, we need the following lemma about simplicity of eigenvalues. It is illuminating to check explicitly how the assumption $v^{\mathrm{T}} u \neq 0$ below fails for a 2×2 Jordan block.

Lemma 1.2.8 (Algebraic simplicity from geometric simplicity)**.** *Let $\lambda \in \mathbb{C}$ be a geometrically simple eigenvalue of some $A \in \mathbb{C}^{n\times n}$. If there exist eigenvectors u and v of A and A^{T} respectively for the eigenvalue λ satisfying $v^{\mathrm{T}} u \neq 0$, then λ is even an algebraically simple eigenvalue of A.*

Proof. Let $x \in \mathbb{C}^n$. It suffices to show that if $(\lambda - A)^2 x = 0$, then $(\lambda - A)x = 0$, so assume that $(\lambda - A)^2 x = 0$. Since $(\lambda - A)x$ is in the eigenspace $\ker(\lambda - A)$ which is spanned by u, there exists a scalar $\alpha \in \mathbb{C}$ such that $(\lambda - A)x = \alpha u$. Hence,

$$\alpha v^{\mathrm{T}} u = v^{\mathrm{T}}(\lambda - A)x = \left((\lambda - A^{\mathrm{T}})\, v\right)^{\mathrm{T}} x = 0.$$

As $v^{\mathrm{T}} u \neq 0$, this implies that $\alpha = 0$, so $(\lambda - A)x = 0$, as claimed. □

Now we have all the tools that we need to prove the Perron–Frobenius theorem.

Proof of Theorem 1.2.5. (a) We first consider the case $\sigma(A) = \{0\}$. In this case, one has $\mathrm{r}(A) = 0 \in \sigma(A)$. Moreover, there exists an integer $k \geq 1$ such that $A^k = 0$. Choose any non-zero vector $y \in \mathbb{R}^n_+$ and let $j \in \{0, 1, \dots, k-1\}$ be the maximal number for which $x := A^j y \neq 0$. Then x is positive since A^j is positive, and $x \in \ker A$.

Now we consider the more interesting case where $\sigma(A) \neq \{0\}$ and hence $\mathrm{r}(A) > 0$.

Choose an eigenvalue λ of A with modulus $|\lambda| = \mathrm{r}(A)$ and let $z \in \mathbb{C}^n$ be an eigenvector of norm 1 corresponding to λ. For every $s > 1$ one has $\mathcal{R}(s\lambda, A)z = \frac{1}{s\lambda - \lambda} z$, and thus

$$\begin{aligned}\frac{1}{(s-1)\,\mathrm{r}(A)}\,|z| = \left|\frac{1}{s\lambda - \lambda} z\right| = |\mathcal{R}(s\lambda, A)z| = \left|\sum_{k=0}^{\infty} \frac{A^k}{(s\lambda)^{k+1}} z\right| \\ \leq \sum_{k=0}^{\infty} \frac{|A^k|}{|s\lambda|^{k+1}}\,|z| = \sum_{k=0}^{\infty} \frac{A^k}{(s\,\mathrm{r}(A))^{k+1}}\,|z| = \mathcal{R}(s\,\mathrm{r}(A), A)\,|z|;\end{aligned}$$

where the penultimate equality uses the positivity of A^k (Proposition 1.1.5). Here we have twice used the Neumann series representation of the resolvent (Proposition 1.2.7(b)), which is applicable because $|s\lambda|, |s\,\mathrm{r}(A)| > \mathrm{r}(A)$.

If we take norms in the inequality $\frac{1}{(s-1)\,\mathrm{r}(A)}\,|z| \leq \mathcal{R}(s\,\mathrm{r}(A), A)\,|z|$ that we just proved, we get $\frac{1}{(s-1)\,\mathrm{r}(A)} \leq \|\mathcal{R}(s\,\mathrm{r}(A), A)\|$ (see the properties of the norm in Remark 1.1.6), so $\|\mathcal{R}(s\,\mathrm{r}(A), A)\| \to \infty$ as $s \downarrow 1$. By continuity of the resolvent (Proposition 1.2.7(a)), it follows that $\mathrm{r}(A)$ is not in the resolvent set and is thus an eigenvalue of A.

It remains to show the existence of an eigenvector $x \in \mathbb{R}^n_+$ for the eigenvalue $\mathrm{r}(A)$. Consider any sequence (s_k) in $(1, \infty)$ that converges to 1; for each index k we define

$$\alpha_k := \|\mathcal{R}(s_k\,\mathrm{r}(A), A)\,|z|\| \qquad \text{and} \qquad x_k := \frac{\mathcal{R}(s_k\,\mathrm{r}(A), A)\,|z|}{\alpha_k}.$$

We have already seen that $\alpha_k \to \infty$ as $k \to \infty$ and that $x_k \geq 0$ for each k. Moreover,

$$(A - \mathrm{r}(A))x_k = (A - s_k\,\mathrm{r}(A))x_k + (s_k\,\mathrm{r}(A) - \mathrm{r}(A))x_k$$

$$= -\frac{|z|}{\alpha_k} + (s_k - 1)\,\mathrm{r}(A)x_k \to 0.$$

Since $\|x_k\| = 1$ for all k and as the unit sphere in $\mathbb{C}^n$ is compact, there exists a subsequence (x_{k_j}) of (x_k) that converges to a non-zero vector $x \in \mathbb{R}^n_+$. Thus, $(A - \mathrm{r}(A))x = \lim_{k\to\infty}(A - \mathrm{r}(A))x_k = 0$, and so x is an eigenvector of A for the eigenvalue $\mathrm{r}(A)$.

(b) Assume now that all diagonal entries of A are non-zero. Then we can find a number $\delta > 0$ such that $A - \delta \geq 0$. Consider the spectral radius $r := \mathrm{r}(A - \delta)$ of $A - \delta$. Since $A - \delta$ is positive, we can apply (a) to this matrix and thus see that r is an eigenvalue of $A - \delta$ and so $r + \delta$ is an eigenvalue of A. In particular, $0 < r + \delta \leq \mathrm{r}(A)$.

On the other hand, as all eigenvalues of $A - \delta$ are contained in the closed disk $B_{\leq r}(0)$ with radius r and centre 0, so all eigenvalues of A are contained in the disk $B_{\leq r}(\delta)$ with radius r and centre δ. Therefore, $\mathrm{r}(A) \leq r + \delta$. It follows that $\mathrm{r}(A) = r + \delta$. But the circle with radius $r + \delta$ and centre 0 intersects the disk $B_{\leq r}(\delta)$ only in the point $r + \delta$, so A has no further eigenvalue of modulus $r + \delta = \mathrm{r}(A)$.

(c) Finally, assume that A is strongly positive. It follows from (b) that $\mathrm{r}(A) > 0$. Alternatively, we can also see this directly: for every eigenvector $x \in \mathbb{R}^n_+$ of A corresponding to the eigenvalue $\mathrm{r}(A)$ – which exists according to (a) – one has $\mathrm{r}(A)x = Ax \succeq \mathbb{1}$. Hence $\mathrm{r}(A) > 0$ and $x \succeq \mathbb{1}$.

Next we show that the eigenvalue $\mathrm{r}(A)$ is geometrically simple. To this end, let $x \succeq \mathbb{1}$ be an eigenvector of A for the eigenvalue $\mathrm{r}(A)$ and let $y \in \mathbb{R}^n$ be any other eigenvector for the same eigenvalue. Then there exists a number $\gamma \in \mathbb{R} \setminus \{0\}$ such that $x - \gamma y$ is positive, but has at least one component that is 0. If $x - \gamma y$ were non-zero, it would be an eigenvector of A for the eigenvalue $\mathrm{r}(A)$, which would imply $x - \gamma y \succeq \mathbb{1}$, as we have just seen. Thus, $x - \gamma y = 0$, so y is a multiple of x. This proves the geometric simplicity of the eigenvalue $\mathrm{r}(A)$.

To see that $\mathrm{r}(A)$ is algebraically simple, we now use Lemma 1.2.8. By applying (a) to the transposed matrix A^{T}, one gets an eigenvector $y \geq 0$ of A^{T} for the eigenvalue $\mathrm{r}(A^{\mathrm{T}}) = \mathrm{r}(A)$. As $y \neq 0$ and $x \succeq \mathbb{1}$, one has $y^{\mathrm{T}}x > 0$, so Lemma 1.2.8 is applicable and shows that the geometric simplicity of $\mathrm{r}(A)$ implies the algebraic simplicity. □

1.3 Positive matrix semigroups

The powers A^k of a square matrix give the solutions $x\colon \mathbb{N}_0 \to \mathbb{C}^n$ to the difference equation $x(k) = Ax(k-1)$ for $k \in \mathbb{N}$. As in the scalar case, it is natural to study the continuous time analogue of this dynamical system, i.e. the differential equation $\dot{x}(t) = Ax(t)$ with $x\colon [0,\infty) \to \mathbb{C}^n$. For this, one uses the matrix exponential function.

Definition 1.3.1 (Matrix exponential function)**.** For every $A \in \mathbb{C}^{n\times n}$ one defines

$$e^A := \exp(A) := \sum_{k=0}^{\infty} \frac{A^k}{k!} \in \mathbb{C}^{n\times n},$$

where the series converges absolutely in $\mathbb{C}^{n\times n}$.

We first discuss a number of essential properties of the matrix exponential function, in particular its relation to linear differential equations. Positivity takes the stage back in Theorems 1.3.8 and 1.3.9.

Proposition 1.3.2 (Properties of the matrix exponential function)**.** *The matrix exponential function has the following properties:*

(a) $e^0 = \mathrm{id}$.

(b) *The matrix exponential function* $\exp\colon \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$, $A \mapsto e^A$ *is continuous.*

(c) *For fixed* $A \in \mathbb{C}^{n\times n}$*, the mapping* $\mathbb{C} \to \mathbb{C}^{n\times n}$, $z \mapsto e^{zA}$ *is differentiable, and hence analytic, with derivative* $\frac{\mathrm{d}}{\mathrm{d}z} e^{zA} = Ae^{zA} = e^{zA}A$ *at each* $z \in \mathbb{C}$.

(d) *If two matrices* $A, B \in \mathbb{C}^{n\times n}$ *satisfy* $AB = BA$*, then* $e^{A+B} = e^A e^B$.

Proof. (a) This follows readily from the definition of the matrix exponential function.

(b) Let $A, B \in \mathbb{C}^{n\times n}$. An induction argument yields the geometric sum formula

$$A^k - B^k = \sum_{j=0}^{k-1} A^j (A-B) B^{k-1-j}$$

for all integers $k \geq 1$. On the right hand side, it is important to have $A - B$ in the middle since A and B are not assumed to commute. Thus we can estimate $\left\|A^k - B^k\right\| \leq k\alpha^{k-1} \|A - B\|$ with $\alpha := \max\{\|A\|, \|B\|\}$. The continuity now follows from

$$\left\|e^A - e^B\right\| \leq \sum_{k=1}^{\infty} \frac{1}{k!} \left\|A^k - B^k\right\| \leq \|A - B\| \sum_{k=1}^{\infty} \frac{\alpha^{k-1}}{(k-1)!} = e^{\alpha} \|A - B\|.$$

(c) This can be shown in the same way as for the scalar-valued exponential function.

(d) One can prove this by using the Cauchy product formula for infinite series, as in the scalar-valued case. Readers familiar with the uniqueness theorem for ordinary differential equations might also find the following alternative proof insightful.

Consider the functions $X_1, X_2\colon \mathbb{R} \to \mathbb{C}^{n\times n}$ that are given by

$$X_1(t) = e^{t(A+B)} \qquad \text{and} \qquad X_2(t) = e^{tA}e^{tB}$$

for all $t \in \mathbb{R}$. According to (a) and (c) the function X_1 solves the initial value problem

$$\begin{cases} \dot{X}(t) = (A+B)X(t) & \text{for all } t \in \mathbb{R}, \\ X(0) = \mathrm{id}. \end{cases}$$

On the other hand, as A and B commute, the definition of the matrix exponential function implies that B also commutes with e^{tA} for all $t \in \mathbb{R}$. This together with (c) and the product rule for differentiation implies that X_2 solves the same initial value problem. Hence, by the uniqueness theorem for linear initial value problems it follows that $X_1(t) = X_2(t)$ for all $t \in \mathbb{R}$. For $t = 1$ this gives the claim. □

Proposition 1.3.2(c) has the following consequence, which is the main reason why one is interested in matrix exponential functions.

Corollary 1.3.3 (Solutions to linear differential equations)**.** *Let $A \in \mathbb{C}^{n\times n}$ and $x_0 \in \mathbb{C}^n$. Then the function $x : [0,\infty) \to \mathbb{C}^{n\times n}$, $x \mapsto e^{tA}x_0$ satisfies the initial value problem*

$$\begin{cases} \dot{x}(t) = Ax(t) & \text{for all } t \in [0,\infty), \\ x(0) = x_0. \end{cases}$$

From the uniqueness theorem for ordinary differential equations, the function x is in fact the only solution to the initial value problem in Corollary 1.3.3.

For a matrix $A \in \mathbb{C}^{n\times n}$, Corollary 1.3.3 shows that the matrix family $(e^{tA})_{t\geq 0}$ is a quite fundamental object. Hence, it gets its own name, which is inspired by the property $e^{(s+t)A} = e^{sA}e^{tA}$ for all $s,t \geq 0$ that follows from Proposition 1.3.2(d).

Definition 1.3.4 (Matrix semigroups and positivity)**.**

(a) Let $A \in \mathbb{C}^{n\times n}$. The family $(e^{tA})_{t\geq 0}$ is called the **matrix semigroup** generated by A.

(b) Let $A \in \mathbb{R}^{n\times n}$. Then $(e^{tA})_{t\geq 0}$ is called **positive** if $e^{tA} \geq 0$ for all $t \in [0,\infty)$.

We have seen (in Proposition 1.2.2) that the spectral radius of a matrix A determines the long-term behaviour of the powers A^k. For the matrix semigroup $(e^{tA})_{t\geq 0}$, a similar role is played by the so-called **spectral bound**.

Definition 1.3.5 (The spectral bound of a matrix)**.** Let $A \in \mathbb{C}^{n\times n}$. The number

$$\mathrm{s}(A) := \max\{\operatorname{Re}\lambda : \lambda \in \sigma(A)\}$$

is called the **spectral bound** of A.

Proposition 1.3.6 (Convergence to 0 of matrix semigroups)**.** *For each matrix $A \in \mathbb{C}^{n\times n}$, the following are equivalent:*

(i) $\mathrm{s}(A) < 0$.

(ii) $e^{tA} \to 0$ *as* $t \to \infty$.

(iii) *There exist numbers $\mu < 0$ and $c \geq 0$ such that $\|e^{tA}\| \leq ce^{t\mu}$ for each $t \geq 0$.*

Proof. *"(i) ⇒ (iii)":* As in the proof of Proposition 1.2.2, this can be deduced using the Jordan normal form of A. We refer to Exercise 1.4(d) for a detailed discussion; cf. proof of Proposition 1.2.2.

"(iii) ⇒ (ii)": This implication is obvious.

"(ii) ⇒ (i)": Let λ be an eigenvalue of A with real part $\operatorname{Re}\lambda = \mathrm{s}(A)$ and an associated eigenvector $z \in \mathbb{C}^n$ of norm one. For every $k \in \mathbb{N}_0$ one has $A^k z = \lambda^k z$ and thus, $e^{tA}z = e^{t\lambda}z$ for every $t \geq 0$ by the definition of the matrix exponential function. So $e^{t\,\mathrm{s}(A)} = e^{t\operatorname{Re}\lambda}\|z\| = \|e^{tA}z\| \to 0$ for each $t \to \infty$, which shows that $\mathrm{s}(A) < 0$. □

The Neumann series representation of the resolvent of a matrix A (given in Proposition 1.2.7(b)) has the following analogue in continuous time.

Lemma 1.3.7 (Laplace transform representation of the resolvent)**.** *Let $A \in \mathbb{C}^{n\times n}$. For every $\lambda \in \mathbb{C}$ that satisfies $\operatorname{Re}\lambda > \operatorname{s}(A)$ one has*

$$\mathcal{R}(\lambda, A) = \int_0^\infty e^{-t\lambda} e^{tA} \, \mathrm{d}t,$$

where the integral converges absolutely.

Proof. Let $\operatorname{Re}\lambda > \operatorname{s}(A)$. Then $\operatorname{s}(A-\lambda) < 0$ and so by Proposition 1.3.6, there are numbers $\mu < 0$ and $c \geq 0$ such that $\left\| e^{-t\lambda} e^{tA} \right\| \leq c e^{t\mu}$ for all $t \geq 0$. Hence, the integral indeed converges absolutely.

To prove that the integral equals $\mathcal{R}(\lambda, A)$, observe that

$$(\lambda - A) \int_0^\infty e^{-t\lambda} e^{tA} \, \mathrm{d}t = \lim_{T\to\infty} -\int_0^T \frac{\mathrm{d}}{\mathrm{d}t} e^{t(A-\lambda)} \, \mathrm{d}t = \lim_{T\to\infty} \big(-e^{T(A-\lambda)} + \mathrm{id} \big) = \mathrm{id};$$

where the last equality uses again that $\operatorname{s}(A-\lambda) < 0$, which indeed gives $e^{T(A-\lambda)} \to 0$ as $T \to \infty$ according to Proposition 1.3.6. □

Except in special cases in small dimensions, it is typically not possible to explicitly compute e^{tA} for a given matrix A. Fortunately, one can check positivity of the semigroup $(e^{tA})_{t\geq 0}$ purely in terms of A, as condition (iv) in the following theorem shows.

Theorem 1.3.8 (Characterisation of positive matrix semigroups)**.** *Let $A \in \mathbb{R}^{n\times n}$. The following are equivalent:*

(i) $e^{tA} \geq 0$ *for all real numbers* $t \geq 0$.

(ii) *For all real numbers* $\lambda > \operatorname{s}(A)$ *one has* $\mathcal{R}(\lambda, A) \geq 0$.

(iii) *For all sufficiently large real numbers* $\lambda > \operatorname{s}(A)$ *one has* $\mathcal{R}(\lambda, A) \geq 0$.

(iv) *All off-diagonal entries of A are in $[0, \infty)$, i.e. $A_{jk} \geq 0$ for all indices $j \neq k$.*

Proof. *"(i) ⇒ (ii)":* This follows from the representation of the resolvent $\mathcal{R}(\lambda, A)$ as the Laplace transform of the semigroup $(e^{tA})_{t\geq 0}$ given in Lemma 1.3.7.

"(ii) ⇒ (iii)": This implication is obvious.

"(iii) ⇒ (iv)": Consider numbers $\lambda \in \mathbb{R}$ that satisfy $\lambda > \operatorname{r}(A)$. The Neumann series representation of the resolvent (Proposition 1.2.7(b)) shows that

$$\lambda^2 \mathcal{R}(\lambda, A) - \lambda \, \mathrm{id} = \sum_{k=1}^\infty \frac{A^k}{\lambda^{k-1}} \to A$$

as $\lambda \to \infty$. For indices $j \neq k$ one thus gets

$$A_{jk} = \lim_{\lambda\to\infty} \big(\lambda^2 \mathcal{R}(\lambda, A) - \lambda \, \mathrm{id} \big)_{jk} = \lim_{\lambda\to\infty} \lambda^2 \mathcal{R}(\lambda, A)_{jk} \geq 0,$$

where the last inequality follows from (iii).

"(iv) ⇒ (i)": As (iv) holds there exists a number $c \in \mathbb{R}$ such that $A + c\,\mathrm{id} \geq 0$. So

$$0 \leq e^{t(A+c\,\mathrm{id})} = e^{tc\,\mathrm{id}} e^{tA} = e^{tc} e^{tA}$$

for all $t \in [0, \infty)$, where the inequality at the beginning follows from $A + c\,\mathrm{id} \geq 0$ and the definition of matrix exponential function, and the first equality follows from Proposition 1.3.2(d). Division by the numbers $e^{tc} \in (0, \infty)$ yields (i). □

Several other equivalent conditions for positivity of matrix semigroups can be found in Exercise 1.2. We conclude this lecture with a Perron–Frobenius type theorem for positive matrix semigroups. It is remarkable that in part (b) of the following theorem, no additional assumption on A is needed. This is in sharp contrast to situation for single operators, where we needed an additional assumption in Theorem 1.2.5(b).

Theorem 1.3.9 (Perron–Frobenius for positive matrix semigroups)**.** *Let $A \in \mathbb{R}^{n\times n}$ and assume the matrix semigroup $(e^{tA})_{t\geq 0}$ is positive.*

(a) $\mathrm{s}(A)$ *is an eigenvalue of A and there exists a corresponding eigenvector $x \geq 0$.*

(b) $\mathrm{s}(A)$ *is a* ***strictly dominant*** *eigenvalue of A in the sense that* $\mathrm{Re}\,\lambda < \mathrm{s}(A)$ *for all other eigenvalues λ of A.*

Proof. Since $(e^{tA})_{t\geq 0}$ is positive, there exists $c \in \mathbb{R}$ such that $A + c\,\mathrm{id} \geq 0$ by Theorem 1.3.8. Therefore by the Perron–Frobenius theorem for positive matrices (Theorem 1.2.5), the spectral radius $\mathrm{r}(A + c\,\mathrm{id})$ is an eigenvalue of $A + c\,\mathrm{id}$ with a positive eigenvector. Using the equality

$$\sigma(A + c\,\mathrm{id}) = \sigma(A) + c,$$

we deduce that $\mathrm{r}(A + c\,\mathrm{id}) = \mathrm{s}(A + c\,\mathrm{id}) = \mathrm{s}(A) + c$. Both assertions are now immediate. □

We end this chapter by pointing out that a similar result as in Theorem 1.2.5(c) can also be proved for matrix semigroups if e^{tA} is strongly positive for every $t > 0$. However, this leaves the natural question how to check the strong positivity of e^{tA} in terms of A. Since this is a bit tangential to the subsequent chapters, we refrain from pursuing this topic further.

Exercises for Chapter 1

Exercise 1.1. Let $A, B \in \mathbb{C}^{n\times n}$.

(i) Give an example to show that $e^{A+B} = e^A e^B$ does not imply $AB = BA$.

(ii) If there exists $\varepsilon > 0$ such that $e^{t(A+B)} = e^{tA}e^{tB}$ for all $t \in [0, \varepsilon)$, then show $AB = BA$.

Exercise 1.2 (Continuation of Theorem 1.3.8)**.** Let $A \in \mathbb{R}^{n\times n}$ be given. Prove that the following are equivalent:

(iv) All off-diagonal entries of A are in $[0,\infty)$, i.e. $A_{jk} \geq 0$ for all indices $j \neq k$.

(v) The matrix A satisfies the *positive minimum principle*, i.e. for all $u \in \mathbb{R}^n_+$ and all $k \in \{1,\dots,n\}$ with $u_k = 0$ one has $(Au)_k \geq 0$.

(vi) The matrix A is *cross positive*, i.e. for all $u, v \in \mathbb{R}^n_+$ with $u^\mathrm{T} v = 0$ one has $u^\mathrm{T} A v \geq 0$.

(vii) The matrix A satisfies the *Beurling–Deny criterion*, i.e. for every $u \in \mathbb{R}^n$ one has $(u^-)^\mathrm{T} A u^+ \geq 0$, where

$$(u^+)_k := \begin{cases} u_k & \text{if } u_k \geq 0, \\ 0 & \text{if } u_k < 0 \end{cases}$$

for all $k \in \{1,\dots,n\}$, and where $u^- := (-u)^+$.

(viii) The matrix A satisfies the *Arendt–Kato inequality*, i.e. for all $u \in \mathbb{R}^n$ and all indices k with $u_k \geq 0$ one has $(Au^+)_k \geq (Au)_k$.

Exercise 1.3. The koala (*Phascolarctos cinereus*) is a notoriously lazy animal, sleeping up to 20 hours a day. It is also a very picky eater. Suppose that a particular koala has 3 favourite eucalyptus trees, arranged as in Figure 1.3.1.

For $i, j \in \{1,2,3\}$, let P_{ij} denote the probability that the koala will eat at tree i the following day given that it has eaten at tree j today. Consider the following model:

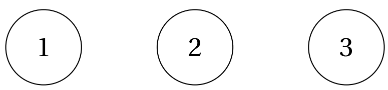

Figure 1.3.1: Eucalyptus trees in an Australian forest (some imagination is required).

- With probability $q \in (0,1)$, the koala will stay at the same tree the following day.
- Since it is lazy, the koala will only move to adjacent trees. Hence, if it has eaten at tree 2 on one day, it will move to either tree 1 or 3 the next day with equal probability (or otherwise stay in place). On the other hand, if it has eaten at tree 1 or 3, it will only move to tree 2 (or otherwise stay in place).

(a) For the probabilities P_{ij} described above, write down the matrix $P = (P_{ij})_{1\le i,j\le 3}$, which is called the **transition matrix** of the model, in terms of q. Explain what the (i,j)-th entry of the matrix powers P^k represents.

(b) Show that $\mathrm{r}(P) = 1$ and that 1 is a strictly dominant eigenvalue.

(c) In the long run, what can you say about the proportion of days the koala spends at each tree?

Exercise 1.4. Let $\lambda_0 \in \mathbb{C}$ and let $J_0 \in \mathbb{C}^{n_0 \times n_0}$ denote the Jordan block

$$J_0 = \begin{pmatrix} \lambda_0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0 \end{pmatrix}.$$

(a) Find and prove a formula for the matrix J_0^k for every $k \in \mathbb{N}_0$.

(b) Find and prove a formula for the matrix e^{tJ_0} for every $t \in [0,\infty)$.

Hint: First consider the case $\lambda_0 = 0$ and then use Proposition 1.3.2(d).

(c) Let $A \in \mathbb{C}^{n\times n}$ such that $\mathrm{r}(A) > 0$. Use the Jordan normal form of A to show that there exists a number $c > 0$ such that $\|A^k\| \le c\,\mathrm{r}(A)^k(1 + k^{n-1})$ for each $k \in \mathbb{N}_0$.

(d) Let $A \in \mathbb{C}^{n\times n}$. Use the Jordan normal form of A to show that there exists a number $c > 0$ such that $\|e^{tA}\| \le c e^{t\,\mathrm{s}(A)}(1 + t^{n-1})$ for each $t \in [0,\infty)$.

Notes for Chapter 1

Positivity versus non-negativity

As promised in Remark 1.1.2, we now discuss the terminology *positive* and the related question of whether 0 is considered positive in a bit more detail.

Real numbers: It is remarkable that even for real numbers, the meaning of the term *positive* depends on the language. The convention that positivity of a real number α means $\alpha > 0$ – while the property $\alpha \geq 0$ is often referred to as α being *non-negative* – is common in English and, for instance, also in German. On the other hand, in French the adjective *positif* typically refers to a number $\alpha \geq 0$.

As the real numbers are defined as an ordered field with a number of additional properties, it is worthwhile to also take a brief look at conventions in the theory of ordered groups and fields. Unsurprisingly, the French meaning of 'positive' is employed by Bourbaki in its definition of ordered groups [Bou07, p. A VI.4]. The same convention is then used in the English translation [Bou03, p. A VI.4].

Linear algebra: A substantial amount of literature studies order properties of $\mathbb{R}^n$ and matrices, in particular in relation to the Perron–Frobenius theorem and its applications. In this field, it seems to be most common to call a vector x *positive* if $x_k > 0$ for all indices k (note that we call this property *strongly positive* in Definition 1.2.3). Vectors $x \geq 0$ are typically referred to as *non-negative* vectors in this part of the literature. This terminology has the advantage that it is consistent with the standard conventions for real numbers in English.

Care must be taken, though, since the coordinate-wise relation $\leq$ defines only a partial order on $\mathbb{R}^n$ when $n \geq 2$: there exist vectors $x \in \mathbb{R}^n$ that satisfy neither $x \geq 0$ nor $x \leq 0$. If one gives in to the temptation to call a vector $x \in \mathbb{R}^n$ *negative* if $-x$ is positive, then the common terminology in linear algebra leads to the following situation: x is negative if and only if $x_k < 0$ for all indices k and thus we have a linguistically unpleasant situation where the assertion "x is non-negative" is inequivalent to "x is not negative".

Real-valued functions: Some parts of the literature adopt a convention similar to the one in linear algebra: a function $f\colon \Omega \to \mathbb{R}$ defined on a set Ω is called *non-negative* if $f(\omega) \geq 0$ for all $\omega \in \Omega$ or, for instance in the setting of L^p-spaces, for almost all $\omega \in \Omega$. Accordingly, f is then called *positive* if $f(\omega) > 0$ for (almost) all $\omega \in \Omega$.

This adaptation of the finite-dimensional perspective comes with an additional caveat in infinite dimensions that only becomes apparent when one develops a systematic theory of ordered spaces in infinite dimensions. The property $f(\omega) > 0$ for (almost) all $\omega \in \Omega$ has very different consequences depending on the surrounding space. For example, in $C(K)$ it implies that f dominates a strictly positive constant on K, whereas in $L^p(\Omega)$ it does not. We will elaborate on this later in Chapter 6,when we have enough Banach lattice theory available.

Elements of ordered vector spaces and Banach lattices: In the theory of ordered vector spaces and Banach lattices, which we introduce in Chapter 6, it is common to call a vector x *positive* if $x \geq 0$; in particular, the zero vector is positive. We follow this convention since we will frequently use Banach lattice theory later on. To maintain consistency throughout these notes, we have adopted the same convention in the finite-dimensional setting, as can be seen in Definition 1.1.1.

Perron–Frobenius and friends

The story of the Perron-Frobenius theorem, and the theory of non-negative matrices[2] in general, has a surprising beginning. At the turn of the 20th century, at the University of Munich (LMU), Oskar Perron was studying the problem of convergence of continued fraction algorithms (German: *Kettenbruchalgorithmen*), following the work of his colleague Alfred Pringsheim. His breakthrough was to reduce the problem to a study of the eigenvalue equation for specific matrices with positive entries (although he did not use this terminology). Consequently Perron was able to simplify the convergence criteria from earlier works (such as those of Pringsheim), and moreover, his methods could be extended to treat the more general case of Jacobi algorithms. This became the subject of his *Habilitation* paper [Per07a], published in *Mathematische Annalen* in 1907.

Clearly Perron recognised the utility of his methods beyond their original purpose and the potential for a systematic theory, for he then followed up with the article *Zur Theorie der Matrices*, which was also published in the *Mathematische Annalen* [Per07b]. In this work, Perron's main theorem corresponds more or less to Theorem 1.2.5(c) in this chapter. Moreover, he could derive the same conclusions under the weaker assumption that $A \geq 0$ and A^k is positive (in our terminology, strongly positive) for some $k \in \mathbb{N}$. However, he expressed dissatisfaction with his rather convoluted argument to achieve this generalisation, and in addition he left open the possibility that a larger class of non-negative matrices could satisfy the conclusions of his theorem.

At this point, Frobenius enters the story. In a series of three papers [Fro08, Fro09, Fro12], he manages to resolve the issues raised by Perron. In a 1908 paper and its sequel in 1909, he proves strengthened versions of Perron's results for positive matrices

[2] In this historical account, we use the classical terminology from linear algebra as explained at the beginning of these notes.

using thoroughly linear-algebraic techniques, especially the properties of determinants. Then, in the 1912 paper, Frobenius extends his investigations to encompass irreducible non-negative matrices. Here, *irreducible* refers to matrices that cannot be put into block upper-triangular form via simultaneous row or column permutations. We encourage readers who are interested in further details (both mathematical and biographical) of the history of the Perron-Frobenius theorem to consult the article [Haw08].

The study of non-negative matrices, stemming from the ideas of Perron and Frobenius, has proved to be very fruitful, and has found diverse applications in the natural and social sciences, from population models (e.g. Leslie matrices) to queuing theory, to input-output models in economics (e.g. the Leontief model), to Google's PageRank algorithm. Thus, the literature on 'Perron-Frobenius theory' is vast. Two texts that are now considered classical include [Sen06], which deals with applications to probability theory and in particular Markov chains, and [BP94], which is notable for its systematic study of positivity with respect to general cones in $\mathbb{R}^n$.

Finally, we cannot omit a mention of the fairly recent monograph [BKFR17], which was based on the material of the 17th Internet Seminar (2013–2014). Part I of that book contains an accessible exposition of the Perron-Frobenius theorem (in particular, Frobenius' contributions), properties of (positive) matrix exponential functions, and numerous applications.

Bibliography

[BKFR17] András Bátkai, Marjeta Kramar Fijavž, and Abdelaziz Rhandi. *Positive operator semigroups*, volume 257 of *Operator Theory: Advances and Applications*. Birkhäuser/Springer, Cham, 2017. From finite to infinite dimensions. With a foreword by Rainer Nagel and Ulf Schlotterbeck.

[Bou03] Nicolas Bourbaki. *Elements of mathematics. Algebra II. Chapters 4–7. Transl. from the French by P. M. Cohn and J. Howie.* Berlin: Springer, reprint of the 1990 English translation edition, 2003.

[Bou07] Nicolas Bourbaki. *Éléments de mathématique. Algèbre. Chapitres 4 à 7.* Berlin: Springer, reprint of the 1981 original edition, 2007.

[BP94] Abraham Berman and Robert J. Plemmons. *Nonnegative matrices in the mathematical sciences*, volume 9 of *Class. Appl. Math.* Philadelphia, PA: SIAM, 1994.

[Fro08] Georg Frobenius. Über Matrizen aus positiven Elementen. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 471–476, 1908.

[Fro09] Georg Frobenius. Über Matrizen aus positiven Elementen. II. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 514–518, 1909.

[Fro12] Georg Frobenius. Über Matrizen aus nicht negativen Elementen. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 456–477, 1912.

[Haw08] Thomas Hawkins. Continued fractions and the origins of the Perron-Frobenius theorem. *Arch. Hist. Exact Sci.*, 62(6):655–717, 2008.

[Mac00] Charles R. MacCluer. The many proofs and applications of Perron's theorem. *SIAM Review*, 42(3):487–498, 2000.

[Per07a] Oskar Perron. Grundlagen für eine theorie des jacobischen kettenbruchalgorithmus. *Math. Ann.*, 64(1):1–76, 1907.

[Per07b] Oskar Perron. Zur Theorie der Matrices. *Math. Ann.*, 64(2):248–263, 1907.

[Sen06] Eugene Seneta. *Non-negative matrices and Markov chains.* Springer Ser. Stat. New York, NY: Springer, revised reprint of the 2nd ed. edition, 2006.