Lecture 2: Positivity and the Perron–Frobenius Theorem

Pierre Magal

University of Bordeaux, France

pierre.magal@math.u-bordeaux.fr

The conference material is taken from my book

Differential Equations and Population Dynamics I, Springer 2022

Winter School on Mathematical Modelling in Epidemiology and Medicine 2023,

June 19 - 24, Institute of Complex Systems of Valparaíso, Valparaíso, Chile

Outline

Definition 1.1

Let $A \in M_n(\mathbb{R})$ be an n by n matrix with real elements. Define the *spectral bound*

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

where $\sigma(A)$ is the *spectrum* of A and is defined by

$$\sigma(A) = \{ \lambda \in \mathbb{C} : \lambda I - A \text{ is not invertible} \} .$$

Then the following lemma holds.

Lemma 1.2

For each $\lambda > s(A)$, the matrix $(\lambda I - A)$ is invertible and we have the following formula

$$(\lambda I - A)^{-1} = \int_0^\infty \mathrm{e}^{-\lambda t} \mathrm{e}^{At} \mathrm{d}t,$$

where $(\lambda I - A)^{-1}$ is called the resolvent of A.

Remark 1.3

The integral $\int_0^\infty \mathrm{e}^{-\lambda t}\mathrm{e}^{At}\mathrm{d}t$ *in the resolvent formula is called the Laplace transform of* $t \to \mathrm{e}^{At}$ *. Note that this integral converges since* $s(A-\lambda I) = s(A) - \lambda < 0$ *and the convergence follows from the stability theorem.*

Remark 1.4

The above formula remains true for any $\lambda \in \mathbb{C}$ *such that*

$$\mathrm{Re}(\lambda) > s(A).$$

In that case we have

$$s(A-\lambda I) = s(A) - \mathrm{Re}\,(\lambda) < 0.$$

Proof. Let $\lambda > s(A)$ be given. By the stability theorem in Chapter 3, the integral $\int_0^\infty \mathrm{e}^{-\lambda t}\mathrm{e}^{At}\mathrm{d}t$ converges. We have

$$(\lambda I - A)\int_0^\infty \mathrm{e}^{-\lambda t}\mathrm{e}^{At}\mathrm{d}t = -\int_0^\infty (-\lambda I + A)\mathrm{e}^{(-\lambda I + A)t}\mathrm{d}t = -\left[\mathrm{e}^{(-\lambda I + A)t}\right]_0^\infty,$$

which yields

$$(\lambda I - A)\int_0^\infty \mathrm{e}^{-\lambda t}\mathrm{e}^{At}\mathrm{d}t = \int_0^\infty \mathrm{e}^{-\lambda t}\mathrm{e}^{At}\mathrm{d}t(\lambda I - A) = I.$$

This ends the proof of the lemma. ☐

By using the Laplace transform we also prove the following lemma.

Lemma 1.5

Assume that 0 is exponentially asymptotically stable for the system

$$X'(t) = AX(t).$$

Then $s(A) < 0$.

Proof. Let $\delta > 0$ and $M > 1$ be such that

$$\|\mathrm{e}^{At}\| \leq M\mathrm{e}^{-\delta t}, \ \forall t \geq 0.$$

Let $\lambda \in \mathbb{C}$ with $\mathrm{Re}(\lambda) > -\delta$. Then the integral

$$\int_0^{\infty} \mathrm{e}^{-\lambda t}\mathrm{e}^{At}\mathrm{d}t$$

converges and

$$(\lambda I - A)\int_0^{\infty} \mathrm{e}^{-\lambda t}\mathrm{e}^{At}\mathrm{d}t = \int_0^{\infty} \mathrm{e}^{-\lambda t}\mathrm{e}^{At}\mathrm{d}t(\lambda I - A) = I.$$

Therefore the matrix $(\lambda I - A)$ is invertible and $\lambda \notin \sigma(A)$. □

Outline

In this section, we introduce partial orders on $\mathbb{R}^n$ and $M_n(\mathbb{R})$.

Definition 2.1

Let $x, y \in \mathbb{R}^n$ be given. We will use the following notations

$$\begin{aligned} x \geq y &\Leftrightarrow x_i \geq y_i \text{ for all } i = 1, \cdots, n, \\ x > y &\Leftrightarrow x \geq y \text{ and } x_{i_0} > y_{i_0} \text{ for some } i_0 \in \{1, \cdots, n\}, \\ x \gg y &\Leftrightarrow x_i > y_i \text{ for all } i = 1, \cdots, n. \end{aligned}$$

Remark 2.2

We can reformulate the above definitions by considering the difference $x - y$ *as follows. We have*

$$\begin{aligned} x \geq y \ \Leftrightarrow \ x-y \geq 0 \ \Leftrightarrow \ x-y \in \mathbb{R}^n_+, \\ x > y \Leftrightarrow \ x-y > 0 \ \Leftrightarrow \ x-y \in \mathbb{R}^n_+ \setminus \{0\}, \\ x \gg y \ \Leftrightarrow \ x-y \gg 0 \ \Leftrightarrow \ x-y \in \mathrm{Int}\left(\mathbb{R}^n_+\right) = (0,\infty)^n. \end{aligned}$$

Remark 2.3

We can generalize the above notion of partial order by replacing $\mathbb{R}^n_+$ *by* $K \subset \mathbb{R}^n$*, a positive cone, that is,* K *is a subset with the following properties:*

(i) K *is closed and convex;*

(ii) $\mathbb{R}_+ K \subset K, \quad$ *where* $\mathbb{R}_+ K := \{\lambda\, x : \lambda \in \mathbb{R}_+$ *and* $x \in K\}$*;*

(iii) $K \cap (-K) = \{0\}$.

Then one can define the partial order $\leq_K$ *by*

$$x \leq_K y \;\Leftrightarrow\; y - x \in K.$$

Similarly to the elements of $\mathbb{R}^n$ we can define a partial order on the space of real matrices $M_n(\mathbb{R})$ as follows: for $A = (a_{i,j})_{i,j=1,\cdots,n} \in M_n(\mathbb{R})$, we set

$$\begin{aligned} A \geq 0 \ &\Leftrightarrow \ a_{ij} \geq 0, \ \forall i, j = 1, \cdots, n, \\ A > 0 \ &\Leftrightarrow \ A \geq 0 \text{ and } A \neq 0, \\ A \gg 0 \ &\Leftrightarrow \ a_{ij} > 0, \ \forall i, j = 1, \cdots, n. \end{aligned}$$

Definition 2.4

Let $x \in \mathbb{C}^n$ and $A \in M_n(\mathbb{C})$ be given. We define the *modulus* of x and A respectively by

$$|x| = \begin{pmatrix} |x_1| \\ |x_2| \\ \vdots \\ |x_n| \end{pmatrix} \in \mathbb{R}_+^n \text{ and } |A| = \begin{pmatrix} |a_{11}| & \dots & |a_{1n}| \\ \vdots & & \vdots \\ |a_{n1}| & \dots & |a_{nn}| \end{pmatrix} \in M_n(\mathbb{R}_+).$$

Outline

Consider the linear Cauchy problem

$$u'(t) = Au(t),\ \forall t \geq 0 \text{ and } u(0) = u_0 \in \mathbb{R}^n.$$

Definition 3.1

We will say that the system is *positivity-preserving* if for any non-negative initial distribution $u_0 \geq 0$ the corresponding solution $t \to u(t)$ stays positive for all time $t \geq 0$. That is,

$$u_0 \geq 0 \ \Rightarrow \ u(t) \geq 0,\ \forall t \geq 0.$$

Remark 3.2

It is clear that the system is positivity-preserving if and only if

$$\mathrm{e}^{At} \geq 0, \quad \forall t \geq 0.$$

The main result of this section is the following theorem.

Theorem 3.3

The two following properties are equivalent

(i) *The system is positivity-preserving.*

(ii) *The off-diagonal elements of A are all non-negative. That is,*

$$a_{ij} \geq 0, \text{ whenever } i, j = 1, \cdots, n \text{ and } i \neq j.$$

Remark 3.4

The above property (ii) *is equivalent to the existence of a real number* $\lambda \geq 0$ *such that*

$$(A + \lambda I) \geq 0.$$

Proof. ($\Rightarrow$) Assume that

$$e^{At} \geq 0, \ \forall t \geq 0.$$

Then for all $\lambda \in \mathbb{R}$ we also have

$$e^{-\lambda t} e^{At} = e^{(A-\lambda I)t} \geq 0, \ \forall t \geq 0.$$

Next, using the resolvent formula stated in Lemma 1.2, we deduce that

$$(\lambda I - A)^{-1} = \int_0^{\infty} e^{(A-\lambda I)t} dt \geq 0, \ \forall \lambda > s(A).$$

Furthermore, for all $\lambda > \max\left(s(A), \|A\|_{\mathcal{L}(\mathbb{R}^n)}\right)$ one has

$$\begin{aligned}(\lambda I - A)^{-1} &= \lambda^{-1}\left(I - \lambda^{-1} A\right)^{-1} \\ &= \lambda^{-1} \sum_{k=0}^{\infty} \frac{A^k}{\lambda^k} = \lambda^{-1}\left(I + \frac{A}{\lambda} + \cdots + \frac{A^k}{\lambda^k} + \cdots\right).\end{aligned}$$

Hence we get

$$\lambda^2 (\lambda I - A)^{-1} = \left(\lambda I + A + \frac{\lambda A^2}{\lambda^2}\left(I - \frac{A}{\lambda}\right)^{-1}\right).$$

Let $\{e_1, \cdots, e_n\}$ denote the canonical basis of $\mathbb{R}^n$. Let $i, j \in \{1, \dots, n\}$ with $i \neq j$. Then we have $\langle e_i, e_j \rangle = 0$ and it follows that

$$0 \leq \langle e_i, \lambda^2 (\lambda I - A)^{-1} e_j \rangle = \langle e_i, A e_j \rangle + \left\langle e_i, \frac{\lambda A^2}{\lambda^2} \left(I - \frac{A}{\lambda} \right)^{-1} e_j \right\rangle,$$

and

$$\left\langle e_i, \frac{A^2}{\lambda} \left(I - \frac{A}{\lambda} \right)^{-1} e_j \right\rangle = O\left(\frac{1}{\lambda} \right) \to 0, \text{ as } \lambda \to \infty.$$

Therefore

$$a_{ij} = \langle e_i, A e_j \rangle \geq 0,$$

and the first implication follows.

($\Leftarrow$) Conversely, assume that all the off-diagonal elements of A are non-negative. Let $\delta > 0$ such that

$$(A + \delta I) \geq 0.$$

Then one has

$$e^{\delta t} e^{At} = \sum_{k=0}^{\infty} \frac{(A + \delta I)^k t^k}{k!} \geq 0, \ \forall t \geq 0,$$

and the result follows. $\square$

Outline

The theorem presented in this chapter was proved by Perron [20] and Frobenius [7, 8]. This is a classical result in matrix analysis. We refer to the books of Gantmacher [9], Seneta [22], Minc [18] and Horn and Johnson [12] for more results on non-negative matrices.

Definitions and notations for matrices

Recall that the linear operator norm of an n by n matrix A with real elements, denoted by $\|A\|_{\mathcal{L}(\mathbb{R}^n)}$, is defined as

$$\|A\|_{\mathcal{L}(\mathbb{R}^n)} := \sup_{\|x\|\leq 1} \|Ax\|.$$

By using this norm we have the following result.

Lemma 4.1

Let $A \in M_n(\mathbb{R})$ *be given. The sequence* $\left(\|A^p\|_{\mathcal{L}(\mathbb{R}^n)}^{1/p}\right)_{p\geq 1}$ *converges as* $p \to \infty$.

The proof of this result given below is taken from the book of Kato [13, pp. 27–28].

Proof. Set $a_p := \ln(\|A^p\|_{\mathcal{L}(\mathbb{R}^n)})$ for $p \geq 0$. Here $A^0 := I$, which implies that $\|A^0\| = 1$, so that $a_0 = 0$. Let us prove that

$$\lim_{p \to +\infty} \frac{a_p}{p} = \inf_{m>0} \frac{a_m}{m}.$$

Since we have

$$\|A^{p+q}\|_{\mathcal{L}(\mathbb{R}^n)} \leq \|A^p\|_{\mathcal{L}(\mathbb{R}^n)} \|A^q\|_{\mathcal{L}(\mathbb{R}^n)}, \ \forall p, q \in \mathbb{N},$$

we deduce that the sequence $(a_p)_{p\geq 0}$ is sub-additive, namely it satisfies

$$a_{p+q} \leq a_p + a_q \ \forall p, q \in \mathbb{N}.$$

Let $m > 0$ be a fixed integer. Applying this last inequality to $p = m \times q + r$, for some $q \in \mathbb{N}$ and $r \in \{0, ..., m-1\}$, we have

$$a_p \leq a_{mq} + a_r \leq q a_m + a_r.$$

Hence for all p large enough (such that $q > 0$) we have

$$\frac{a_p}{p} \leq \frac{q}{p} a_m + \frac{1}{p} a_r \leq \frac{1}{m + \frac{r}{q}} a_m + \frac{1}{p} a_r$$

and since $q \to +\infty$ whenever $p \to +\infty$, we obtain

$$\limsup_{p \to +\infty} \frac{a_p}{p} \leq \frac{a_m}{m}.$$

Now since this inequality must be true for all $m > 0$ we deduce that

$$\limsup_{p \to +\infty} \frac{a_p}{p} \leq \inf_{m>0} \frac{a_m}{m}.$$

Now for each integer $p > 0$ we have

$$\frac{a_p}{p} \geq \inf_{m>0} \frac{a_m}{m},$$

so that we get

$$\liminf_{p \to +\infty} \frac{a_p}{p} \geq \inf_{m>0} \frac{a_m}{m} \geq \limsup_{p \to +\infty} \frac{a_p}{p}.$$

The result follows. □

Definition 4.2

The *spectral radius* $r(A)$ of a matrix $A \in M_n(\mathbb{R})$ is defined by

$$r(A) := \lim_{p \to +\infty} \|A^p\|_{\mathcal{L}(\mathbb{R}^n)}^{\frac{1}{p}}.$$

By using the Jordan normal form of A we deduce the following result.

Lemma 4.3

The following equality holds

$$r(A) = \sup\{|\lambda| : \ \lambda \in \sigma(A)\}.$$

Exercise 4.4

Prove the above lemma (Hint: Use Jordan's reduction of A*).*

By using the same arguments as for the above discrete time case we have the following result.

Lemma 4.5

Let $A \in M_n(\mathbb{R})$ *be given. The limit* $\lim_{t \to +\infty} \frac{1}{t} \ln \|\mathrm{e}^{tA}\|_{\mathcal{L}(\mathbb{R}^n)}$ *exists.*

Definition 4.6

The *growth rate* of the semigroup $\{\mathrm{e}^{At}\}_{t\geq 0}$ is defined as

$$\mathrm{Gr}(A) := \lim_{t\to+\infty} \frac{1}{t} \ln \|\mathrm{e}^{tA}\|_{\mathcal{L}(\mathbb{R}^n)}.$$

By using the Jordan normal form of A again we also deduce the following result.

Lemma 4.7

The growth rate and the growth bound coincide. That is,

$$\mathrm{Gr}(A) = \sup\{\mathrm{Re}\,(\lambda) : \ \lambda \in \sigma(A)\} =: s(A).$$

Figure 1 summarizes the above notions for the spectrum.

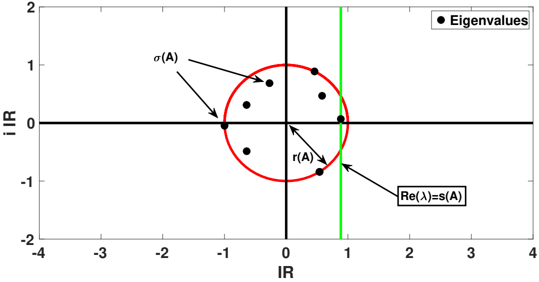


Figure: *In this figure the black dots represent the point of the spectrum in the complex plane. In general, the spectral radius does not belong to the spectrum of* A*. The growth bound* $s(A)$ *corresponds to the green line. The peripheral spectrum is the spectrum of* A *(*$i.e.$ *the black dots) belonging to the red circle. Observe that, in general, the growth bound (green line) is strictly smaller than the spectral radius* $r(A)$*. The Perron–Frobenius theorem implies that the growth bound and the spectral radius are equal for non-negative matrices (because for non-negative matrices* $r(A)$ *belongs to the spectrum of* A*).*

Definition 4.8

The *peripheral spectrum* of A is defined as

$$\sigma_{\text{per}}(A) = \{\lambda \in \sigma(A) : \ |\lambda| = r(A)\}.$$

Definition 4.9

The *algebraic multiplicity* of an eigenvalue λ_0 of A is the dimension of the generalized eigenspace of A associated to the eigenvalue λ_0 and is also its multiplicity as a root of the characteristic polynomial of A, i.e. the algebraic multiplicity of λ_0 is the integer $n_0 \geq 1$ satisfying

$$\det(\lambda I - A) = (\lambda - \lambda_0)^{n_0} g(\lambda),$$

where $\lambda \to g(\lambda)$ is a polynomial satisfying $g(\lambda_0) \neq 0$. The *geometric multiplicity* of an eigenvalue λ_0 of A is the dimension of the eigenspace $N(\lambda_0 I - A)$.

Remark 4.10

By considering the Jordan normal form of a given matrix we can prove that the algebraic multiplicity is always greater than or equal to the geometric multiplicity. More precisely, the geometric multiplicity is equal to the number of Jordan blocks, while the algebraic multiplicity is the number of times that the eigenvalue appears on the diagonal of the Jordan reduction.

Definition 4.11

We will say that an eigenvalue of A is *simple* if its algebraic multiplicity is 1.

Remark 4.12

An eigenvalue λ_0 *is simple if and only if*

$$\dim(N(\lambda_0 I - A)) = 1$$

and

$$N(\lambda_0 I - A) = N(\lambda_0 I - A)^2.$$

The Perron–Frobenius theorem for strictly positive matrices

Theorem 4.13

Let $A \in M_n(\mathbb{R})$ be a strictly positive matrix (i.e. $A \gg 0$). Then A satisfies the following properties

- (i) $r(A) > 0$.
- (ii) $r(A)$ *is an eigenvalue of* A.
- (iii) $r(A)$ *is the unique eigenvalue of* A *with modulus* $r(A)$ *or equivalently* $\sigma_{per}(A) = \{r(A)\}$.
- (iv) *There exists a* $v_r \gg 0$ *(a right eigenvector of* A*) and there exists a* $v_l \gg 0$ *(a left eigenvector of* A*) such that*

$$Av_r = r(A)v_r \text{ and } v_l^T A = r(A)v_l^T.$$

- (v) $r(A)$ *is a simple eigenvalue of* A.

Example. In order to illustrate this theorem we consider the transition matrix of a *homogeneous Markov chain* (see Seneta [22] for more results on this topic). Namely consider a random variable with a finite number of states $1, \ldots, n$. Assume that the system is in the i^{th} state at time t and the next jump will take it to the j^{th} state (at time $t+1$) with probability m_{ij}. Therefore

$$P[X = j | X = i] = m_{ij}.$$

Then it holds that

$$\sum_{j=1}^{n} m_{ij} = 1, \quad \forall i = 1, \cdots, n.$$

Define

$$M = (m_{ij})_{i,j=1,\dots,n}.$$

Then, one has

$$M \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

As an example with $n = 4$ we consider the following matrix

$$M = \frac{1}{10} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \\ 3 & 4 & 1 & 2 \\ 2 & 3 & 4 & 1 \end{pmatrix}.$$

The matrix M is said to be *Markovian* because

$$M \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Now by the Perron–Frobenius theorem for strictly positive matrices, there exists a vector $V \gg 0$ such that

$$V^T M = r(M) V^T.$$

So we obtain

$$V^T \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = V^T M \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = r(M) V^T \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

This yields

$$V^T \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = r(M) V^T \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix},$$

and since $V \gg 0$, one deduces that $r(M) = 1$.

Let Π^t denote the row vector of the probability distribution of X_t at time t. Then it satisfies

$$\Pi^{t+1} = \Pi^t M, \quad \forall t \geq 0.$$

Then whenever M is primitive we have

$$\lim_{t \to \infty} \frac{1}{\Pi_1^t + \cdots + \Pi_n^t} \Pi^t = \frac{1}{V_1 + \cdots + V_n} V,$$

and since $\Pi_1^t + \cdots + \Pi_n^t = 1$ this is also equivalent to

$$\lim_{t \to \infty} \Pi^t = \frac{1}{V_1 + \cdots + V_n} V.$$

Remark 4.14

For the example $M = \frac{1}{10}\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \\ 3 & 4 & 1 & 2 \\ 2 & 3 & 4 & 1 \end{pmatrix}$ *the left eigenvector is also* $V = (1, 1, 1, 1)^T$.

In Figure 2 we plot the eigenvalues of the matrix M in the complex plane.

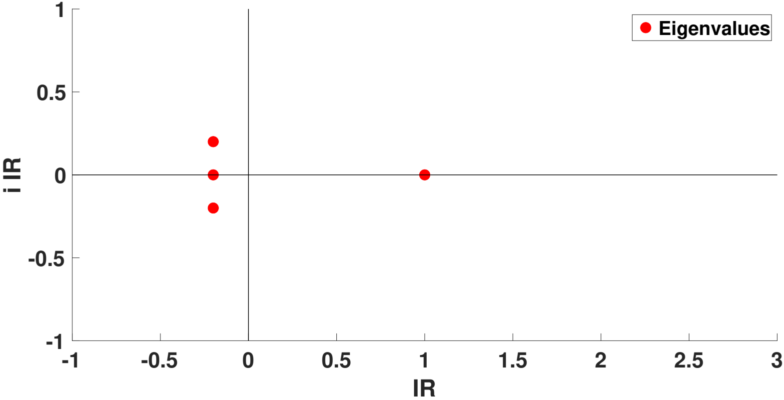


Figure: *In the figure we plot the eigenvalues of M in the complex plane. We observe that two eigenvalues are real (the spectral radius 1 and a negative eigenvalue -0.2) and the two last eigenvalues are complex conjugates $-0.2 \pm 0.2i$.*

End example.

Lemma 4.15

Let $z_1, z_2, \ldots, z_n \in \mathbb{C} \setminus \{0\}$ *be given, for some* $n \geq 2$. *Then*

$$\left| \sum_{k=1}^{n} z_k \right| = \sum_{k=1}^{n} |z_k|$$

if and only if for each $j \in \{2, \ldots, n\}$, *there exists an* $\alpha_j > 0$ *such that*

$$z_j = \alpha_j z_1.$$

Proof. The proof of the above lemma is based on an induction argument. For $n = 2$, note that using polar coordinates one can rewrite z_1 and z_2 as $z_1 = r_1 \mathrm{e}^{i\theta_1}$ and $z_2 = r_2 \mathrm{e}^{i\theta_2}$. Hence we get

$$\begin{aligned}
&|z_1 + z_2| = |z_1| + |z_2| \\
&\Leftrightarrow |r_1 \mathrm{e}^{i\theta_1} + r_2 \mathrm{e}^{i\theta_2}| = r_1 + r_2 \\
&\Leftrightarrow |r_1 + r_2 \mathrm{e}^{i(\theta_2 - \theta_1)}| = r_1 + r_2 \\
&\Leftrightarrow |r_1 + r_2 \cos((\theta_2 - \theta_1)) + i r_2 \sin((\theta_2 - \theta_1))| = r_1 + r_2 \\
&\Leftrightarrow (r_1 + r_2 \cos((\theta_2 - \theta_1)))^2 + (r_2 \sin((\theta_2 - \theta_1)))^2 = (r_1 + r_2)^2
\end{aligned}$$

and after simplification this leads us to

$$2 r_1 r_2 \cos((\theta_2 - \theta_1)) = 2 r_1 r_2.$$

Now since by assumption we have $z_1 \neq 0$ and $z_2 \neq 0$, we end up with $\cos((\theta_2 - \theta_1)) = 1$, that is $(\theta_2 - \theta_1) = 2k\pi$, for some integer $k \in \mathbb{Z}$. Hence we deduce that the result holds true for $n = 2$.

Assume that the result holds true for each integer $2 \leq n \leq n_0$, for some integer $n_0 \geq 2$ and let us prove that the property also holds true for $n_0 + 1$. So assume that

$$|z_1 + z_2 + \cdots + z_{n_0+1}| = |z_1| + |z_2| + \cdots + |z_{n_0+1}|. \tag{1}$$

By using the triangle inequality we have

$$|z_1 + \cdots + z_{n_0} + z_{n_0+1}| \leq |z_1 + \cdots + z_{n_0}| + |z_{n_0+1}|.$$

Next by using (1) and again the triangle inequality we deduce that

$$|z_1| + \cdots + |z_{n_0}| \leq |z_1 + \cdots + z_{n_0}| \leq |z_1| + \cdots + |z_{n_0}|.$$

This implies that

$$|z_1| + \cdots + |z_{n_0}| = |z_1 + \cdots + z_{n_0}|.$$

By using our induction assumption we obtain that for each $j \in \{2, \ldots, n_0\}$, there exists an $\alpha_j > 0$ such that

$$z_j = \alpha_j z_1,$$

and setting $C := 1 + \alpha_2 + \ldots + \alpha_{n_0} > 0$ the equality (1) becomes

$$|Cz_1 + z_{n_0+1}| = |Cz_1| + |z_{n_0+1}|.$$

The result follows by using the case $n = 2$. This completes the proof of the lemma. $\square$

Proof of Theorem 4.13. *Proof of (i).* The trace of A is given by

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i,$$

where the $\lambda_i \in \mathbb{C}$ are eigenvalues of A (listed according to their algebraic multiplicities).

Since $a_{ii} > 0$ for all $i = 1, \cdots, n$ we obtain that $\mathrm{tr}(A) > 0$, which implies that $r(A) > 0$.

Indeed if $r(A) = 0$, then $\lambda_i = 0$ for all $i = 1, \cdots, n$, and $\mathrm{tr}(A) = 0$, which is a contradiction.

Proof of (ii). Since $r(A) > 0$ by the previous step, multiplying A by $r(A)^{-1}$ we can assume that

$$r(A) = 1.$$

Moreover, due to Lemma 4.3, there exists at least one eigenvalue $\lambda \in \mathbb{C}$ of A such that

$$|\lambda| = r(A) = 1.$$

Let $x \neq 0$ be an eigenvector associated to λ. Recall that we have defined the modulus of a matrix $|M|$ as the matrix with elements $|m_{ij}|$. With this notation we have

$$|x| = |\lambda||x| = |\lambda x| = |Ax| \leq |A||x|.$$

Since the matrix A is positive, we have $A = |A|$ and we obtain

$$|x| \leq A|x|.$$

Let us prove that $A|x| = |x|$. To see this, assume that

$$A|x| > |x| \Leftrightarrow A|x| - |x| > 0.$$

Since $A \gg 0$, we deduce that

$$A(A|x| - |x|) = A^2|x| - A|x| \gg 0.$$

So there exists an $\varepsilon > 0$ such that

$$A^2|x| - A|x| > \varepsilon A|x| \Leftrightarrow \frac{1}{1+\varepsilon}A^2|x| > A|x|.$$

By setting

$$B := \frac{1}{1+\varepsilon} A,$$

the above inequality can be rewritten as

$$BA|x| > A|x|.$$

By applying the positive matrix B on both sides of this inequality we obtain

$$B^2A|x| > BA|x| > A|x|,$$

and by induction we obtain for each integer $k > 1$,

$$B^kA|x| > B^{k-1}A|x| > \cdots > BA|x| > A|x|.$$

Hence for each $k > 1$ one has

$$B^kA|x| > A|x|. \tag{2}$$

On the other hand we also have

$$r(B) = \frac{1}{1+\varepsilon} r(A) = \frac{1}{1+\varepsilon} < 1,$$

which implies that

$$\lim_{k \to \infty} B^k = 0.$$

Taking the limit as k goes to $+\infty$ in (2), we obtain

$$0 \geq A|x|,$$

which is impossible since $A \gg 0$ and $|x| > 0$ imply that $A|x| \gg 0$. Therefore we deduce that

$$A|x| = |x|,$$

and 1 is an eigenvalue of A.

Proof of (iii). We still assume that $r(A) = 1$. Let λ be an eigenvalue with modulus 1. Let $x \in \mathbb{C}^n \setminus \{0\}$ be an eigenvector of A associated to λ. Then from the previous part of the proof we have

$$A|x| = |x|. \tag{3}$$

This equality implies first that $|x| \gg 0$ (since $A \gg 0$ and $|x| > 0$). Therefore

$$x_j \neq 0, \quad \forall j \in \{1, \ldots, n\}.$$

By using again $A|x| = |x|$, we have for each $j \in \{1, \ldots, n\}$

$$|x_j| = \sum_{k=1}^{n} a_{jk}|x_k| = \sum_{k=1}^{n} |a_{jk}x_k|$$

and since $Ax = \lambda x$ we also have

$$|x_j| = |\lambda||x_j| = |(\lambda x)_j| = |(Ax)_j| = \left|\sum_{k=1}^{n} a_{jk}x_k\right|.$$

Coupling the above two equalities we obtain, for each $j \in \{1, \ldots, n\}$,

$$\sum_{k=1}^{n} |a_{jk}x_k| = \left|\sum_{k=1}^{n} a_{jk}x_k\right|.$$

Let $j \in \{1,\dots,n\}$ be fixed and set

$$z_k := a_{jk}x_k \in \mathbb{C}, \quad \forall k \in \{1,\dots,n\}.$$

Then the above equality can be rewritten as

$$\left|\sum_{k=1}^{n} z_k\right| = \sum_{k=1}^{n} |z_k| .$$

Next Lemma 4.15 applies and ensures that for each $k \in \{2, \dots, n\}$, there exists an $\alpha_k > 0$ such that

$$z_j = \alpha_j z_1.$$

Hence we can find positive real numbers $\alpha_2 > 0, \dots, \alpha_n > 0$ such that for each $k \in \{2, \dots, n\}$,

$$a_{jk} x_k = \alpha_k a_{j1} x_1,$$

so that the vector x can be rewritten as

$$x = x_1 \begin{pmatrix} 1 \\ \alpha_2 \frac{a_{j1}}{a_{j2}} \\ \vdots \\ \alpha_n \frac{a_{j1}}{a_{jn}} \end{pmatrix}.$$

But since $Ax = \lambda x$, by dividing both sides of this equality by x_1 we obtain

$$\lambda \begin{pmatrix} 1 \\ \alpha_2 \frac{a_{j1}}{a_{j2}} \\ \vdots \\ \alpha_n \frac{a_{j1}}{a_{jn}} \end{pmatrix} = A \begin{pmatrix} 1 \\ \alpha_2 \frac{a_{j1}}{a_{j2}} \\ \vdots \\ \alpha_n \frac{a_{j1}}{a_{jn}} \end{pmatrix}.$$

Thus $\lambda \in (0, \infty)$ and since $|\lambda| = 1$, one obtains $\lambda = 1$. As a consequence 1 is the only eigenvalue with modulus 1.

Proof of (iv). Here we still assume that $r(A) = 1$.

Let $x \in \mathbb{C}^n \backslash \{0\}$ be an eigenvector of A associated to the eigenvalue 1.

Then from the proof of (ii) we have

$$A|x| = |x| \text{ and } |x| \gg 0$$

since $A \gg 0$ and $|x| > 0$.

Therefore one can choose $v_r = |x|$.

Similarly since $r(A) = r(A^T)$ we can apply the same argument to the transposed matrix A^T and we can find $v_l \gg 0$ such that $A^T v_l = v_l$.

Proof of (v). Let us first prove that the geometric multiplicity of $r(A) = 1$ is 1.

Assume by contradiction that

$$\dim(\mathrm{N}(A - r(A)I)) > 1.$$

Then we can find two linearly independent (non-null) vectors $u \in \mathbb{R}^n \setminus \{0\}$ and $v \in \mathbb{R}^n \setminus \{0\}$ such that

$$Au = u \text{ and } Av = v.$$

From part (ii), we have $A|u| = |u|$ and $A|v| = |v|$, which implies that $|u| \gg 0$ $(\Leftrightarrow u_i \neq 0, \forall i = 1, \ldots, n)$ and $|v| \gg 0$ $(\Leftrightarrow v_i \neq 0, \forall i = 1, \ldots, n)$.

Define

$$w := v_1\, u - u_1\, v.$$

Then by construction $w_1 = 0$. Moreover, since $u_1 \neq 0$ and $v_1 \neq 0$ and u and v are linearly independent, we must have

$$w \neq 0,$$

and since u and v are two eigenvectors associated to $r(A) = 1$, we must have

$$Aw = w.$$

We deduce that

$$A|w| = |w| \gg 0,$$

and we obtain a contradiction with the fact that $w_1 = 0$.

We deduce that the geometric multiplicity of $r(A) = 1$ is 1, that is,

$$\dim(\mathrm{N}(A - I)) = 1.$$

Next, assume by contradiction that the algebraic multiplicity of $r(A) = 1$ is not 1. This implies that

$$\mathrm{N}(A - I)^2 \neq \mathrm{N}(A - I).$$

Remember that $\mathrm{N}(A - I) \subset \mathrm{N}(A - I)^2$. It follows that we can find $x \in \mathrm{N}(A - I)^2$ such that

$$x \notin \mathrm{N}(A - I).$$

Therefore we can find

$$Ax = x + y \text{ with } Ay = y \text{ and } y \gg 0.$$

But we also have for each integer $k \in \mathbb{N}$

$$A(x + ky) = (x + ky) + y.$$

Choosing k large enough, one can assume that $z := x + ky \gg 0$ and

$$Az = z + y.$$

We obtain $z \ll Az$ and $z \gg 0$, which is impossible from the proof of part (ii). This completes the proof of the theorem. □

Consider the linear operator $\Pi : \mathbb{R}^n \to \mathbb{R}^n$ defined by

$$\Pi x := \frac{\langle v_l, x \rangle}{\langle v_l, v_r \rangle} v_r.$$

Then Π is a projector, that is,

$$\Pi^2 = \Pi.$$

Moreover, Π commutes with A and we have

$$A\Pi = \Pi A = r(A)\Pi.$$

We deduce that

$$A(I - \Pi) = (I - \Pi)A,$$

and we obtain a state space decomposition

$$\mathbb{R}^n = \mathrm{R}(\Pi) \oplus \mathrm{R}(I - \Pi),$$

where $\mathrm{R}(\Pi)$ (respectively $\mathrm{R}(I - \Pi)$) is the range of Π (respectively $I - \Pi$). Moreover, we have

$$A(\mathrm{R}(\Pi)) \subset \mathrm{R}(\Pi) \text{ and } A(\mathrm{R}(I - \Pi)) \subset \mathrm{R}(I - \Pi).$$

From the Perron–Frobenius theorem we know that $r(A)$ is an eigenvalue of A with algebraic multiplicity 1, so we have

$$\sigma(A|_{\mathrm{R}(I-\Pi)}) = \sigma(A)\backslash\{r(A)\}.$$

Furthermore, since $r(A)$ is the only eigenvalue in the peripheral spectrum of A, we deduce that

$$r(A|_{\mathrm{R}(I-\Pi)}) < r(A).$$

Therefore by choosing $\delta \in \left(\frac{r(A|_{\mathrm{R}(I-\Pi)})}{r(A)}, 1\right)$, one can find a constant $M = M(\delta) \geq 1$ such that

$$\left\|\frac{1}{r(A)^k} A^k (I-\Pi)\right\| \leq M\delta^k, \forall k \in \mathbb{N}.$$

As a consequence of the Perron–Frobenius theorem, we obtain the following.

Corollary 4.16

Let $A \in M_n(\mathbb{R})$ *be a strictly positive matrix (i.e.* $A \gg 0$*). Then the rank* 1 *projector* $\Pi \in M_n(\mathbb{R})$ *defined by*

$$\Pi x := \frac{\langle v_l, x \rangle}{\langle v_l, v_r \rangle} v_r$$

satisfies

$$A\Pi = \Pi A = r(A)\Pi \text{ and } \lim_{k \to +\infty} \frac{1}{r(A)^k} A^k = \Pi.$$

More precisely, we can find two constants $M > 1$ *and* $\delta \in (0,1)$ *such that for any* $x \in \mathbb{R}^n$ *one has*

$$\left\| \Pi x - \frac{1}{r(A)^k} A^k x \right\| = \left\| \frac{1}{r(A)^k} A^k (I - \Pi) x \right\| \leq M \delta^k \| (I - \Pi) x \|.$$

Remark 4.17

The above property is also called the asynchronous exponential growth property. This means that the normalized distribution converges to a distribution that is independent of the initial distribution.

Remark Let $A \in M_n(\mathbb{R})$ be a strictly positive matrix (i.e. $A \gg 0$). Assume that the dynamical distribution of population is described by the difference equation in $\mathbb{R}^n$

$$N(t+1) = AN(t), \quad \forall t \in \mathbb{N} \text{ and } N(0) = N_0 \geq 0.$$

Consider a left eigenvector $v_l \gg 0$ associated to the spectral radius $r(A)$. Then one has

$$\langle v_l, N(t+1) \rangle = \langle v_l, AN(t) \rangle = r(A) \langle v_l, N(t) \rangle$$

and by induction we get

$$\langle v_l, N(t) \rangle = r(A)^t \langle v_l, X_0 \rangle, \ \forall t \in \mathbb{N}.$$

In addition, observe that the map

$$\|x\|_1 := \langle v_l, |x| \rangle, \ \forall x \in \mathbb{R}^n$$

is a norm on $\mathbb{R}^n$.
As a consequence, when a population density is described by such a discrete-time model, the Perron–Frobenius theorem provides an equivalent indicator of the Malthusian growth. Namely, we have

$$\|N(t)\|_1 = r(A)^t \|N_0\|_1, \quad \forall t \in \mathbb{N}.$$

Moreover, for $N_0 > 0$ in $\mathbb{R}^n$, the asynchronous exponential growth as stated in Corollary 4.16 means the following convergence for the normalized distribution

$$\lim_{t \to +\infty} \frac{N(t)}{\|N(t)\|_1} = \frac{\Pi N_0}{\langle v_l, N_0 \rangle} = \frac{v_r}{\langle v_l, v_r \rangle}.$$

Herein $v_r \gg 0$ denotes a right eigenvector associated to the spectral radius $r(A)$.

Primitive and irreducible matrices

In this section, we extend the Perron–Frobenius theorem to a larger class of matrices, the class of so-called primitive and irreducible matrices.

Definition 4.18

Let $A \in M_n(\mathbb{R})$ be a non-negative matrix. We will say that A is *primitive* if there exists an integer $m \geq 1$ such that

$$A^m \gg 0.$$

We will say that a matrix $A \in M_n(\mathbb{R})$ is *irreducible* if there exists an integer $m \geq 1$ such that

$$I + A + A^2 + \cdots + A^m \gg 0.$$

Remark 4.19

A matrix is reducible if it is not irreducible. For a reducible matrix A*, one can find that a permutation of the elements of the basis, such that the matrix of* A *expressed in the permuted basis is block lower triangular*

$$B = \begin{bmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{bmatrix},$$

where B_{11} *and* B_{22} *are both square blocks.*

Definition 4.20

Let $A = (a_{ij})$ be an $n \times n$ non-negative matrix. Consider n distinct points P_1, P_2, ..., P_n in the plane, which we call *nodes*. If $a_{ij} > 0$, we connect node P_j to P_i by means of a directed path. The graph obtained is called the *graph* $G(A)$ *associated with the matrix* A.

Definition 4.21

A directed graph $G(A)$ is *strongly connected* if for any pair of nodes P_i and P_j there exists a directed path connecting P_i to P_j (such a direct path can eventually be composed by several single paths joining some intermediate nodes).

Proposition 4.22

The following properties are equivalent

(i) A *is irreducible.*

(ii) *The matrix* $\varepsilon I + A$ *is primitive for all* $\varepsilon > 0$*.*

(iii) *For each* $i, j \in 1, \ldots, n$*, there exists an integer* $m = m(i,j) > 0$ *such that*

$$\langle e_j, A^m e_i \rangle > 0,$$

where $\{e_1, \ldots, e_n\}$ *denotes the canonical basis of* $\mathbb{R}^n$*.*

(iv) *The directed graph* $G(A)$ *of* A *is strongly connected.*

Proof. *Proof of* $(i) \Leftrightarrow (ii)$. Let $m > 0$ be a given integer and $\varepsilon > 0$. By using the binomial formula we have

$$(\varepsilon I + A)^m = \sum_{k=0,\ldots,m} C_n^k \varepsilon^{n-k} A^k$$

with $A^0 = I$. Next we can find two numbers $0 < c_- < c_+$ such that

$$c_-(I + A + A^2 + \cdots + A^m) \leq (\varepsilon I + A)^m \leq c_+(I + A + A^2 + \cdots + A^m).$$

The proofs of $(i) \Leftrightarrow (iii)$ and $(i) \Leftrightarrow (iv)$ are left as an exercise. □

The following theorem can be found in the book of Horn and Johnson [12, Theorem 8.5.3, p. 517].

Theorem 4.23

A non-negative matrix A is primitive if and only if $G(A)$ is strongly connected and the greatest common divisor of the lengths of all paths from P_i to itself is one.

Remark 4.24

Horn and Johnson [12, Corollary 8.5.9, p. 520] proved by using graph theory applied to $G(A)$ *that the matrix* A *is primitive if and only if*

$$A^{n^2} \gg 0.$$

whenever A *is a* n *by* n *matrix.*

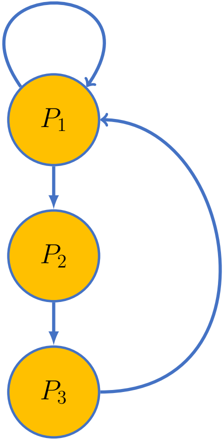
Remark 4.25

Observe that if A *is primitive then* A *is irreducible. The converse is false. Indeed, consider the Leslie matrix* $A = \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right]$. *Then one has*

$$I + A \gg 0.$$

So A *is irreducible. But* A *is not primitive since* $A^{2n} = I, \quad \forall n \in \mathbb{N}$.

Graphs of primitive matrices



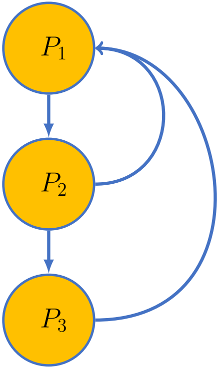


Figure: *In this figure we show some examples of graphs $G(A)$ of primitive, irreducible and non-irreducible matrices.*

Graph of a (non-primitive) irreducible matrix

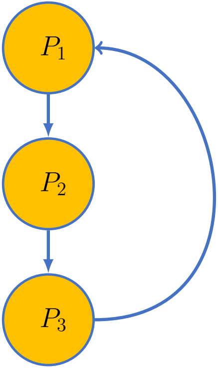


Figure: *In this figure we show some examples of graphs $G(A)$ of primitive, irreducible and non-irreducible matrices.*

Graph of a non-irreducible matrix

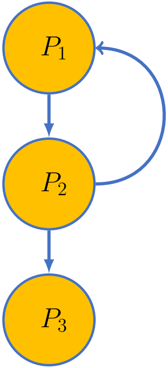


Figure: *In this figure we show some examples of graphs* $G(A)$ *of primitive, irreducible and non-irreducible matrices.*

In Figures above the first primitive graphs correspond, for example, to the matrices

$$\begin{pmatrix} 1/2 & 0 & 1 \\ 1/2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

while the irreducible and non-irreducible graphs correspond to the matrices

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Our next theorem is concerned with primitive matrices.

Theorem 4.26

Let $A \in M_n(\mathbb{R})$ *be a non-negative and primitive matrix. Then the conclusions of Theorem 4.13 and Corollary 4.16 hold.*

Exercise 4.27

Prove Theorem 4.26. Reconsider the arguments given in the proof of Theorem 4.13 for a matrix A *which now is only primitive.*

Hint: By using the Jordan reduction again we can prove that for each integer $n > 0$

$$\sigma(A^n) = \{\lambda^n : \lambda \in \sigma(A)\},$$

and of course one has $|\lambda^n| = |\lambda|^n$*, for all* $\lambda \in \mathbb{C}$*.*

As already mentioned, the matrix

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

is irreducible but not primitive. The characteristic equation is

$$\det(\lambda I - A) = \lambda^2 - 1, \; \forall \lambda \in \mathbb{C}.$$

Therefore the spectrum of A reads as

$$\sigma\left(A\right) = \left\{-1, 1\right\}.$$

So the peripheral spectrum of A is not reduced to its spectral radius $r(A) = 1$.

More generally the same result holds for the following $n \times n$ Leslie matrix

$$L_n = \begin{bmatrix} 0 & 0 & \cdots & 0 & \beta \\ \pi & 0 & \cdots & \cdots & 0 \\ 0 & \pi & 0 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \pi & 0 \end{bmatrix} \in M_n(\mathbb{R}),$$

for some parameters $\beta > 0$ and $\pi > 0$. Note that one has

$$L_n^n = \beta \pi^{n-1} I.$$

So for all eigenvalues $\lambda \in \sigma\left(L\right)$, we have

$$\lambda^n = \beta \pi^{n-1}.$$

Therefore the spectrum and the peripheral spectrum of L coincide.

Our next result is concerned with irreducible matrices.

Theorem 4.28

Let $A \in M_n(\mathbb{R})$ be a non-negative and irreducible matrix. Then the following properties hold:

(i) $r(A) > 0$.

(ii) $r(A)$ *is an eigenvalue of* A.

(iii) *There exists* $v_r \gg 0$ *(a right eigenvector of* A*) and* $v_l \gg 0$ *(a left eigenvector of* A*) such that*

$$Av_r = r(A)v_r \text{ and } v_l^T A = r(A)v_l^T.$$

(iv) $r(A)$ *is a simple eigenvalue of* A.

Moreover, if A *is not primitive, then the peripheral spectrum of* A *contains some eigenvalues distinct from* $r(A)$.

Exercise 4.29

Prove Theorem 4.28.

Hint 1: Let $\varepsilon > 0$ *be given so that the matrix* $A + \varepsilon I$ *is primitive and*

$$\lambda \in \sigma(A) \Leftrightarrow \lambda + \varepsilon \in \sigma(A + \varepsilon I).$$

Hint 2: For the last part of the theorem, assume by contradiction that the peripheral spectrum of A *contains no other eigenvalue than* $r(A)$*. Then as above conclude that*

$$\lim_{k \to +\infty} \frac{1}{r(A)^k} A^k = \Pi \gg 0$$

(where $\Pi = v_r v_l^T$*), which implies that* A *is primitive.*

Example. In Figure 6 we plot the spectrum of

$$L_4 = \begin{bmatrix} 0 & 0 & 0 & \beta \\ \pi & 0 & 0 & 0 \\ 0 & \pi & 0 & 0 \\ 0 & 0 & \pi & 0 \end{bmatrix}.$$

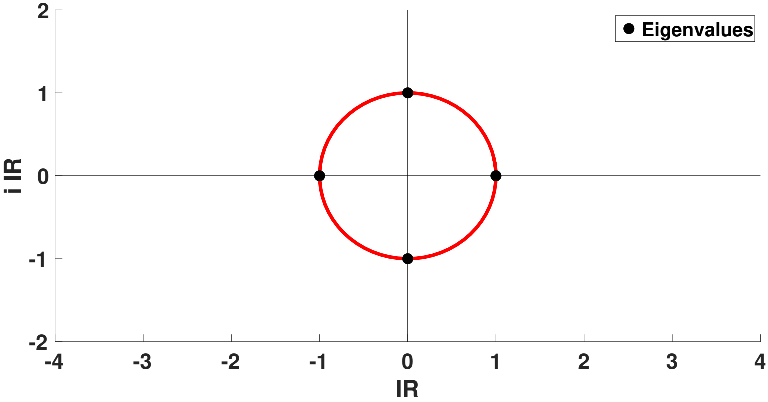


Figure: *In this figure we plot the spectrum of the above Leslie matrix* L_4 *with* $\pi = 1$ *and* $\beta = 1$.

In Figure 7 we plot the spectrum of

$$\widehat{L}_4 = \begin{bmatrix} 0 & 0 & \beta & \beta \\ \pi & 0 & 0 & 0 \\ 0 & \pi & 0 & 0 \\ 0 & 0 & \pi & 0 \end{bmatrix}.$$

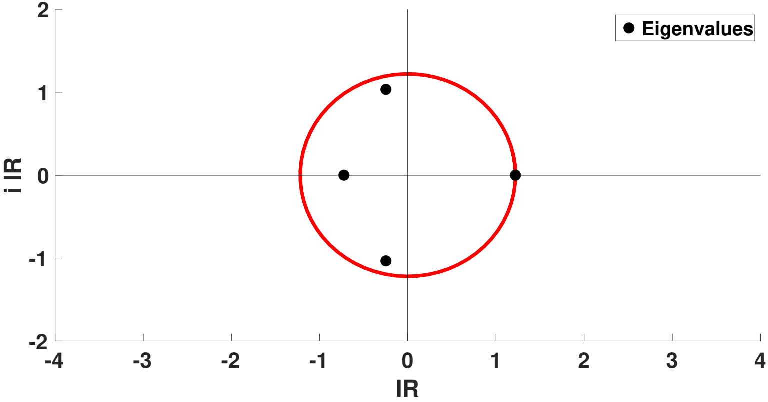


Figure: *In this figure we plot the spectrum of the above Leslie matrix* $\widehat{L}_4$ *with* $\pi = 1$ *and* $\beta = 1$.

The major difference between L_4 and $\widehat{L}_4$ is that $\widehat{L}_4$ is primitive while L_4 is only irreducible.

Outline

Application to Leslie's model

In this section, we reconsider the Leslie matrix introduced in Chapter 1. Consider the Leslie model

$$U^{t+1} = LU^t, \text{ for all } t \geq 0, \text{ with } U^0 = U_0 \in \mathbb{R}_+^{m+1},$$

where the Leslie matrix L is given by

$$L = \begin{bmatrix} \beta_0 & \beta_1 & \cdots & \cdots & \beta_m \\ \pi_0 & 0 & \cdots & \cdots & 0 \\ 0 & \pi_1 & 0 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \pi_{m-1} & 0 \end{bmatrix}.$$

Here we assume that

$$\beta_j \geq 0, \quad \forall j = 0, \ldots, m$$

and

$$\pi_j > 0, \quad \forall j = 0, \ldots, m-1.$$

By computing the powers of the Leslie matrix

$$\widehat{L} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \in M_{m+1}(\mathbb{R}),$$

one obtains the following lemma.

Lemma 5.1

The matrix L is irreducible if $\beta_m > 0$.

The following proposition is also a consequence of Theorem 4.23. This result is due to Demetrius [4].

Proposition 5.2

Assume that $\beta_m > 0$ *and that there exists an integer* $j_0 \in \{0, 1, \dots, m-1\}$ *such that*

$$\beta_{j_0} > 0 \text{ and } \beta_{j_0+1} > 0.$$

Then L *is a primitive matrix.*

Proof. It is also instructive to prove the above proposition by considering the renewal equation for $t > m-1$

$$U_0(t+1) = \beta_0 U_0(t) + \beta_1 \pi_0 U_0(t-1) + \dots + \beta_m \pi_0 \times \pi_{m-1} U_0(t-(m-1)).$$

The proof is left to the reader. ☐

From the above results and the Perron–Frobenius theorem, the following lemma follows.

Theorem 5.3

Assume that $\beta_m > 0$. *Then the spectral radius* $r(L) > 0$ *is the unique positive solution of the* λ*-equation*

$$1 = \frac{1}{\lambda}\left[\beta_0 + \beta_1 \frac{\pi_0}{\lambda} + \beta_2 \frac{\pi_0}{\lambda}\frac{\pi_1}{\lambda} + \cdots + \beta_m \frac{\pi_0}{\lambda}\frac{\pi_1}{\lambda} \times \cdots \times \frac{\pi_{m-1}}{\lambda}\right].$$

Moreover, the right eigenvector of L *associated with* $r(L)$ *is given by*

$$U_r := \begin{pmatrix} 1 \\ \frac{\pi_0}{r(L)} \\ \frac{\pi_0}{r(L)}\frac{\pi_1}{r(L)} \\ \vdots \\ \frac{\pi_0}{r(L)}\frac{\pi_1}{r(L)} \times \cdots \times \frac{\pi_{m-1}}{r(L)} \end{pmatrix}.$$

Theorem 5.4

Furthermore, setting

$$R_0 := \beta_0 + \beta_1\pi_0 + \beta_2\pi_0\pi_1 + \cdots + \beta_m\pi_0\pi_1 \times \cdots \times \pi_{m-1},$$

we have following alternatives:

(i) *If* $R_0 > 1$ *then* $r(L) > 1$.

(ii) *If* $R_0 = 1$ *then* $r(L) = 1$.

(iii) *If* $R_0 < 1$ *then* $r(L) < 1$.

If we assume in addition that L *is primitive, then for each initial distribution* $U_0 > 0$ *the solution of the difference equation*

$$U(t+1) = LU(t), \forall t \geq 0, \text{ and } U(0) = U_0 > 0,$$

satisfies the following asymptotic behavior

$$\lim_{t\to+\infty} \frac{1}{U(t)_0 + U(t)_1 + \ldots + U(t)_m} U(t) = \frac{1}{U_{r0} + U_{r1} + \ldots + U_{rm}} U_r.$$

Figure 8, taken from Chapter 1, illustrates this last convergence result to the right eigenvector.

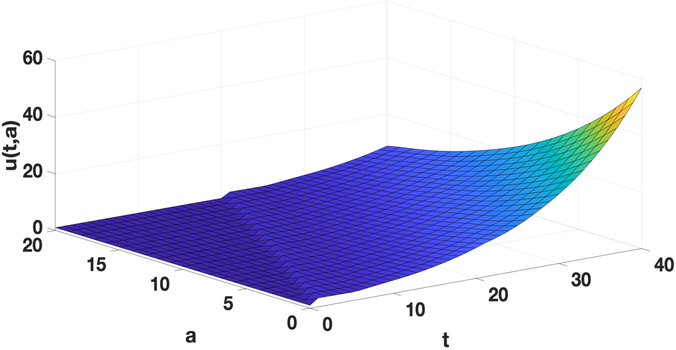


Figure: *In this figure we plot a solution* $t \to u(t,a)$ *of the Leslie model* $a \in [0,20]$*. The reproduction function* $\beta(a) = 0.8 * \Delta a$ *if* $a > 5$ *and* $\beta(a) = 0$ *otherwise. The survival rate is* $\pi(a) = \exp(-0.1 * \Delta a)$*. The initial distribution is constant equal to* 1*. We observe that it takes* 40 *years for the distribution of population to grow exponentially.*

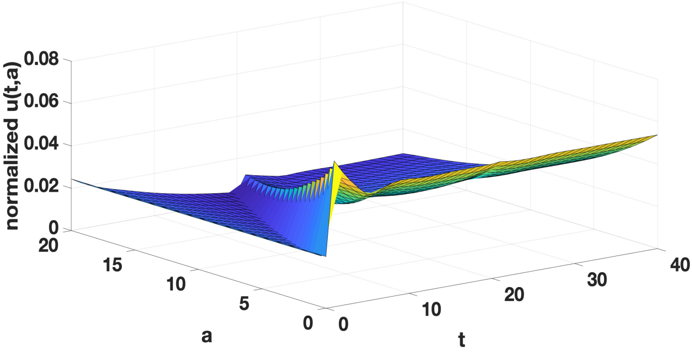


Figure: *In this figure we plot a normalized solution* $t \to u(t,a)/\Sigma_{i=0,\dots,20}u(t,i)$ *of the Leslie model* $a \in [0,20]$. *The reproduction function is defined by* $\beta(a) = 0.8 * \Delta a$ *if* $a > 5$ *and* $\beta(a) = 0$ *otherwise. The survival rate is* $\pi(a) = \exp(-0.1 * \Delta a)$. *The initial distribution is constant equal to* 1. *We observe the convergence of the normalized distribution when the time becomes large enough.*

Note that since $\beta_m > 0$, the matrix L is irreducible (see Lemma 5.1). Hence, due to the Perron–Frobenius theorem (see Theorem 4.23, Theorem 4.28 and its corollary), to prove the above theorem, it is sufficient to check the first assertion on the spectral radius $r(L) > 0$. To see this, note that the equation

$$U \in \mathbb{R}^{m+1} \setminus \{0\} \text{ and } r(L)U = LU,$$

which can be rewritten as $\lambda = r(L) > 0$, satisfies the equation

$$1 = \frac{1}{\lambda}\left[\beta_0 + \beta_1 \frac{\pi_0}{\lambda} + \beta_2 \frac{\pi_0}{\lambda}\frac{\pi_1}{\lambda} + \ldots + \beta_m \frac{\pi_0}{\lambda}\frac{\pi_1}{\lambda} \times \ldots \times \frac{\pi_{m-1}}{\lambda}\right] =: f(\lambda).$$

Next, since $\beta_m > 0$, we have

$$\lim_{\lambda \to 0} f(\lambda) = +\infty \text{ and } \lim_{\lambda \to +\infty} f(\lambda) = 0,$$

and since f is decreasing on $(0, \infty)$, the above equation has a unique solution $\lambda_0 > 0$, that is, $\lambda_0 = r(L)$.

Remark 5.5

Leslie's model was used by Song to understand how to control the growth of the population in China. The number $N_i(t)$ *is the number of females in age class* i. *Assume that* β_i *is the average number of newborns per female in age class* i, *which has been estimate over one year. Assume that* π_i *is the fraction of females surviving from age class* i *to age class* $i+1$. *Then (by using the characteristic equation) one can derive the growth rate* $r(L)$. *Song's work on population control led to the one-child policy in China [24, 25]. More information can be found in Greenhalgh [10, 11].*

More recently, since the growth rate $r(L)$ *has been reduced by the one-child policy, the expected distribution of the population* v_r *has changed. When we normalize* v_r *(i.e. divide by* $v_{r0} + v_{r1} + \cdots + v_{rm}$*) we obtain the expected distribution of population of China. That is, one can compute the expected fraction of young people, middle-aged people, and old people in the population. Since the expected number of old people is now too large compared to the expected number of middle-aged people (or working people), the authorities decided to change the one-child policy, and a second child is now allowed in China.*

Application to the space-time discrete diffusion process

In this section we consider the space-time discrete heat equation with Neumann homogeneous boundary conditions. As discussed in Chapter 1, if $\Delta t > 0$ and $\Delta x > 0$ denote respectively the time and space step, the discrete heat equation reads for $t \geq 0$ as

$$u(t+\Delta t) = u(t) + \frac{\varepsilon \Delta t}{\Delta x^2} D\, u(t) \text{ with } u(0) = u_0 \in \mathbb{R}^N_+.$$

Herein the matrix D is given by

$$D = \begin{pmatrix} -1 & 1 & 0 & \dots & \dots & 0 \\ 1 & -2 & 1 & \ddots & & \vdots \\ 0 & 1 & -2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & 0 \\ \vdots & & \ddots & 1 & -2 & 1 \\ 0 & \dots & & 0 & 1 & -1 \end{pmatrix}.$$

To understand the dynamical behavior of the difference equation, we first investigate some basic properties of the matrix D.

Lemma 5.6

Assume that $p := 2\frac{\varepsilon \Delta t}{\Delta x^2} < 1$. *Then the matrix* $I + \frac{\varepsilon \Delta t}{\Delta x^2} D$ *is non-negative and primitive.*

Remark 5.7

The condition $2\frac{\varepsilon \Delta t}{\Delta x^2} < 1$ *is called the CFL (Courant–Friedrichs–Lax) condition.*

It is readily checked that

$$D\mathbb{1} = 0 \text{ and } \mathbb{1}^T D = 0^T.$$

Hence we get

$$\left(I + \frac{\varepsilon \Delta t}{\Delta x^2} D\right) \mathbb{1} = \mathbb{1} \text{ and } \mathbb{1}^T \left(I + \frac{\varepsilon \Delta t}{\Delta x^2} D\right) = \mathbb{1}^T,$$

and applying the Perron–Frobenius theorem we deduce that the following lemma holds.

Lemma 5.8

Assume that $p = 2\frac{\varepsilon \Delta t}{\Delta x^2} < 1$*. Then the spectral radius of the matrix* $I + \frac{\varepsilon \Delta t}{\Delta x^2} D$ *is given by*

$$r\left(I + \frac{\varepsilon \Delta t}{\Delta x^2} D\right) = 1.$$

Moreover, for each $u_0 \geq 0$ *one has*

$$\sum_{i=1}^{N} u(t)_i = \sum_{i=1}^{N} u_{0i}, \quad \forall t \geq 0,$$

which means that the total number of individuals is constant, and

$$\lim_{n \to +\infty} u(t) = \left(\sum_{i=1}^{N} u_{0i}\right) \begin{pmatrix} 1/N \\ 1/N \\ \vdots \\ 1/N \end{pmatrix},$$

which means that the individuals are equally redistributed between the

Proof. Since by assumption $p < 1$ we deduce by applying Theorem 4.23 that the matrix $M = (I + \frac{\varepsilon \Delta t}{\Delta x^2} D)$ is primitive. Therefore, by using the Perron–Frobenius theorem, we know that there exists a $v_l \gg 0$ such that

$$v_l^T M = r(M) v_l^T.$$

But

$$v_l^T \mathbb{1} = v_l^T M \mathbb{1} = r(M) v_l^T \mathbb{1}.$$

Therefore $r(M) = 1$. The result follows by using the fact that the dimension of the right and left eigenspace of M associated with 1 is one. Therefore the right eigenspace of M associated to 1 is $\mathbb{R}\mathbb{1}$ and the left eigenspace of M associated to 1 is $\mathbb{R}\mathbb{1}^T$. ☐

The Figure below illustrates this convergence result.

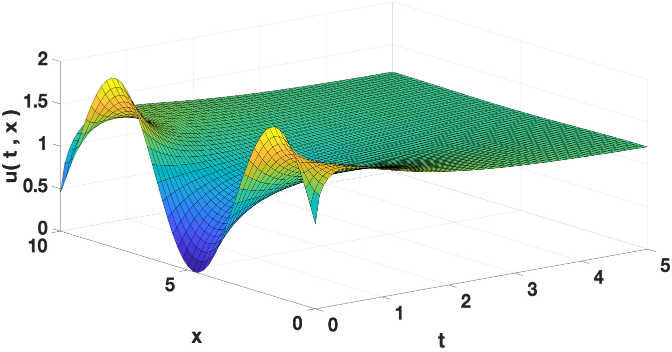


Figure: *In this figure we plot a heat equation with* $x \in [0, 10]$*. The diffusion coefficient is equal to* $\varepsilon = 2$*. The initial distribution is constant equal to* $u_0(x) = 1 + \sin(x)$*. We observe the quite rapid convergence to the constant distribution.*

Exercise 5.9

Assume that $2\frac{\varepsilon\Delta t}{\Delta x^2} < 1$. *Consider the implicit numerical scheme given by*

$$u(t+\Delta t) = u(t) + \frac{\varepsilon\Delta t}{\Delta x^2} D\, u(t+\Delta t),$$

yielding the following linear difference equation

$$\left(I - \frac{\varepsilon\Delta t}{\Delta x^2} D\right) u(t+\Delta t) = u(t), \quad \forall t \geq 0 \text{ and } u(0) = u_0.$$

Prove that if $2\frac{\varepsilon\Delta t}{\Delta x^2} < 1$, *then the matrix* $(I - \frac{\varepsilon\Delta t}{\Delta x^2} D)$ *is invertible and the matrix* $(I - \frac{\varepsilon\Delta t}{\Delta x^2} D)^{-1}$ *is strictly positive.*

The case of linear ordinary differential equations

In this part, we come back to the positivity property of the solutions of linear differential equations and we describe some further properties as a consequence of the Perron–Frobenius theorem. Our result reads as follows.

Theorem 5.10

Let $A \in M_n(\mathbb{R})$ *be a given matrix such that for all* $\delta > 0$ *large enough* $\delta I + A$ *is a non-negative and primitive matrix. Then the following properties hold:*

(i) *For each* $t > 0$ *one has*

$$\mathrm{e}^{At} \gg 0.$$

(ii) *The spectral bound of* A, $\mathrm{s}(A) := \max\{\mathrm{Re}\,(\lambda) : \ \lambda \in \sigma(A)\}$, *is a simple eigenvalue of the matrix* A.

(iii) *For each* $\lambda \in \sigma(A) \backslash \{\mathrm{s}(A)\}$ *one has*

$$\mathrm{Re}\,(\lambda) < \mathrm{s}(A).$$

Theorem 5.11

(iv) *There exist two vectors* $v_r \gg 0$ *and* $v_l \gg 0$ *such that*

$$v_l^T A = \mathrm{s}(A) v_l^T \text{ and } A v_r = \mathrm{s}(A) v_r.$$

(v) *If* Π *is the projector given by* $\Pi(x) := \frac{\langle v_l, x\rangle}{\langle v_l, v_r\rangle} v_r$ *for* $x \in \mathbb{R}^n$, *then we have*

$$\Pi \mathrm{e}^{At} = \mathrm{e}^{At}\Pi = \mathrm{e}^{\mathrm{s}(A)t}\Pi, \quad \forall t \geq 0,$$

and there exist two constants $\chi > 0$ *and* $M \geq 1$ *such that*

$$\|\mathrm{e}^{-s(A)t}\mathrm{e}^{At}(I - \Pi)\| \leq M\mathrm{e}^{-\chi t}\|(I - \Pi)\|, \quad \forall t \geq 0.$$

Remark 5.12

In the above theorem, the diagonal elements of A can have any sign, while all the off-diagonal elements of A are non-negative.

Proof. *Proof of (i).* Let $\delta > 0$ be such that $\delta I + A$ is non-negative and primitive. Then we have

$$\mathrm{e}^{At} = \mathrm{e}^{-\delta t}\mathrm{e}^{(A+\delta I)t} = \mathrm{e}^{-\delta t}\sum_{k=0}^{+\infty}\frac{((A+\delta I)t)^k}{k!} \gg 0, \quad \forall t > 0.$$

Sketch of the proof for (ii)–(iv). To prove (ii) implies (iv), it is sufficient to apply the Perron–Frobenius theorem to the primitive matrix $(A + \delta I)$ and to observe that one has

$$\lambda \in \sigma(A) \Leftrightarrow \lambda + \delta \in \sigma(A + \delta I).$$

Now assertion (v) is a direct consequence of (ii)–(iv). ☐

Example. We reconsider the heat equation, but now we assume the time is continuous while the space is discrete. In other words, we assume that the individuals are located at some discrete positions. Then we can consider the ordinary differential equation

$$u'(t) = \frac{\gamma}{2} D u(t), \forall t \geq 0, \text{ and } u(0) = u_0 \geq 0,$$

where D has been defined in the previous section and $u(t)_i$ is the number of individuals in position i. Then for $i = 2, \dots, N-1$

$$u_i(t)' = \frac{\gamma}{2} \left(u_{i+1} + u_{i-1} \right) - \gamma u_i$$

with

$$u_1(t)' = \frac{\gamma}{2} \left(u_2 - u_1 \right)$$

and

$$u_N(t)' = \frac{\gamma}{2} \left(u_{N-1} - u_N \right).$$

Here the parameter $\gamma > 0$ is the leaving rate of individuals at position i. In other words, the time spent in position i follows the exponential law with average

1

By applying Theorem 5.10 we obtain the following lemma.

Lemma 5.13

For each $u_0 \geq 0$ *we have*

$$\lim_{t \to +\infty} u(t) = (u_{01} + u_{02} + \cdots + u_{0N}) \begin{pmatrix} 1/N \\ 1/N \\ \vdots \\ 1/N \end{pmatrix}.$$

End example.

Stability criteria for linear ordinary differential equations

The stability criteria stated in this part are commonly used in population dynamics. In the context of epidemic models, such criteria have been extensively used in the literature. We refer to Diekmann, Heesterbeek and Metz [5] and to Van den Driessche and Watmough [6]. One can also find some infinite-dimensional versions of this result, for which we refer to Thieme [27].

Example. In order to illustrate this stability result, let us consider the following continuous-time Leslie model

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} U_0(t) \\ U_1(t) \end{pmatrix} = L \begin{pmatrix} U_0(t) \\ U_1(t) \end{pmatrix} \tag{4}$$

wherein the matrix L takes the form

$$L := \begin{pmatrix} \beta_0 - \mu_0 & \beta_1 \\ \eta_0 & -\mu_1 \end{pmatrix}$$

for some given parameters $\beta_0 \geq 0$, $\beta_1 > 0$, $\eta_0 > 0$, $\mu_0 > 0$, $\mu_1 > 0$.
Set

$$A = \begin{pmatrix} -\mu_0 & 0 \\ 0 & -\mu_1 \end{pmatrix} \text{ and } B = \begin{pmatrix} \beta_0 & \beta_1 \\ \eta_0 & 0 \end{pmatrix}.$$

Then we have

$$\begin{aligned} B(-A)^{-1} &= \begin{pmatrix} \beta_0 & \beta_1 \\ \eta_0 & 0 \end{pmatrix} \times \begin{pmatrix} \mu_0^{-1} & 0 \\ 0 & \mu_1^{-1} \end{pmatrix} \\ &= \begin{pmatrix} \beta_0 \mu_0^{-1} & \beta_1 \mu_1^{-1} \\ \eta_0 \mu_0^{-1} & 0 \end{pmatrix}. \end{aligned}$$

The matrix $B(-A)^{-1}$ is irreducible, and by using Lemma 5.4 we deduce that $r(B(-A)^{-1}) < 1$ if

$$R_0 := \beta_0 \mu_0^{-1} + \beta_1 \mu_1^{-1} \eta_0 \mu_0^{-1} < 1.$$

It follows from Theorem 5.14 that (4) is stable (i.e. $s(L) < 0$) if $R_0 < 1$.
End example.

Theorem 5.14 (Stability)

Let $A, B \in M_n(\mathbb{R})$ *be two given matrices with* $\delta I + A \geq 0$ *for some* $\delta \geq 0$ *and* $B \geq 0$*. Assume that* $\mathrm{s}(A) < 0$ *or equivalently the equilibrium* 0 *is exponentially asymptotically stable for the system*

$$X'(t) = AX(t).$$

Assume in addition that the matrix $B(-A)^{-1}$ *is non-negative and irreducible and satisfies*

$$r\left(B(-A)^{-1}\right) < 1.$$

Then one has $\mathrm{s}(A+B) < 0$ *or equivalently the equilibrium* 0 *is exponentially asymptotically stable for the system*

$$X'(t) = (A+B)X(t).$$

Proof. Let us prove that $\mathrm{s}(A+B)<0$. With that aim, let $\lambda \in \mathbb{C}$ be given with $\mathrm{Re}\,(\lambda) \geq 0$. Observe first that since for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ one has

$$(\lambda I-(A+B))x = y \Leftrightarrow [I-B(\lambda I-A)^{-1}](\lambda I-A)x = y,$$

it follows that the matrix $(\lambda I-(A+B))$ is invertible if and only if the matrix $[I-B(\lambda I-A)^{-1}]$ is invertible. Let us now prove that we can find a new norm $\|\cdot\|_1$ on $\mathbb{R}^n$ such that

$$\sup_{\|x\|_1 \leq 1} \|B(\lambda I-A)^{-1}x\|_1 < 1, \quad \forall \lambda \in \{\nu \in \mathbb{C} : \mathrm{Re}\,(\nu) \geq 0\}.$$

Since $B \geq 0$ and $\mathrm{e}^{At} \geq 0$ for all $t \geq 0$, it follows from the resolvent formula and the definition of the modulus that for all $x \in \mathbb{R}^n$ one has

$$|B(\lambda I - A)^{-1}x| = \left| B \int_0^{+\infty} \mathrm{e}^{-\lambda t}\mathrm{e}^{At}x\mathrm{d}t \right| \leq B \int_0^{+\infty} \mathrm{e}^{-\mathrm{Re}\,(\lambda)t}\mathrm{e}^{At}|x|\mathrm{d}t.$$

Now since $\mathrm{Re}\,(\lambda) \geq 0$ we obtain

$$|B(\lambda I - A)x| \leq B \int_0^{+\infty} \mathrm{e}^{At}|x|\mathrm{d}t = B(-A)^{-1}|x|, \quad \forall x \in \mathbb{R}^n.$$

Since the matrix $B(-A)^{-1}$ is irreducible, the Perron–Frobenius theorem applies and ensures that we can find $v_g \gg 0$ such that

$$v_g^T B(-A)^{-1} = r\left(B(-A)^{-1}\right) v_g^T.$$

Define

$$\|x\|_1 := \langle v_g, |x| \rangle = v_g^T |x|, \quad \forall x \in \mathbb{R}^n.$$

Then $\| \cdot \|_1$ is a norm on $\mathbb{R}^n$ and one has

$$\begin{aligned} \|B(\lambda I - A)x\|_1 \quad &= v_g^T |B(\lambda I - A)x| \leq v_g^T B(-A)^{-1} |x| \\ &= r\left(B(-A)^{-1}\right) \|x\|_1, \, \forall x \in \mathbb{R}^n. \end{aligned}$$

Finally, since by assumption one has $r\left(B(-A)^{-1}\right) < 1$, the result follows.
□

Example 5.15

In order to illustrate this stability result, let us consider the following continuous-time Leslie model

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} U_0(t) \\ U_1(t) \\ \vdots \\ U_m(t) \end{pmatrix} = L \begin{pmatrix} U_0(t) \\ U_1(t) \\ \vdots \\ U_m(t) \end{pmatrix} \tag{5}$$

wherein the matrix L takes the form

$$L = \begin{pmatrix} \beta_0 - \mu_0 & \beta_1 & \cdots & \cdots & \beta_m \\ \eta_0 & -\mu_1 & 0 & \cdots & 0 \\ 0 & \eta_1 & -\mu_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \eta_{m-1} & -\mu_m \end{pmatrix}.$$

for some given parameters $\beta_i \geq 0$, $i = 0, \dots, m$, $\beta_m > 0$, $\eta_i > 0$, $\mu_i > 0$,

Outline

Positivity of linear EDO systems in Banach spaces

Definition 6.1

A closed and convex subset X_+ of a Banach space X is called a *positive cone* of X if the following properties are satisfied:

(i) $\lambda X_+ \subset X_+, \forall \lambda \geq 0;$

(ii) $X_+ \cap (-X_+) = \{0\}.$

Similarly as for $\mathbb{R}^n$, by using the positive cone X_+ one can define a partial order on X as follows

$$\begin{aligned} x \geq y \;&\Leftrightarrow\; x-y \in X_+, \\ x > y &\Leftrightarrow\; x-y \in X_+ \setminus \{0\}, \\ x \gg y \;&\Leftrightarrow\; x-y \in \mathring{X}_+, \end{aligned}$$

wherein $\mathring{X}_+$ is the interior of X_+.

Remark 6.2

In a Banach space the interior of the positive cone is not necessarily non-empty. Indeed,

(i) *The interior of the positive cone* $C_+([0,1],\mathbb{R})$ *is not empty in* $C([0,1],\mathbb{R}))$ *(Hint: The function* $u(x) = 1, \forall x \in [0,1]$ *belongs to the interior of the cone);*

(ii) *The interior of the positive cone* $L^1_+((0,1),\mathbb{R})$ *is empty in* $L^1((0,1),\mathbb{R})$ *(Hint: The function* $u(x) = 1,$ *for almost every* $x \in [0,1]$*, does not belong to the interior of the cone).*

Definition 6.3

A Banach space X endowed with such a partial order $\geq$ is said to be *partially ordered*. We will say that a bounded linear operator $A \in \mathcal{L}(X)$ is *positive* (or for short $A \geq 0$) if

$$Ax \geq 0, \ \forall x \geq 0, \text{ that is, } AX_+ \subset X_+.$$

Theorem 6.4

Let $A \in \mathcal{L}(X)$ be a bounded linear operator on a partially ordered Banach space X. Then

$$\mathrm{e}^{At} \geq 0, \quad \forall t \geq 0,$$

if and only if there exists a $\lambda_0 > 0$ large enough such that $[\lambda_0, \infty) \subset \rho(A)$ and

$$(\lambda I - A)^{-1} \geq 0, \quad \forall \lambda \geq \lambda_0.$$

The proof of this theorem is similar to that of Theorem 3.3. This proof uses the fact that the positive cone is closed and we leave it as an exercise. ☐

Further reading about the Perron–Frobenius theorem

The Perron–Frobenius theorem presented in this chapter was proved by Perron [20] and Frobenius [7, 8]. This is one of the most classical results about the spectrum of matrices. We refer to the books of Gantmacher [9], Seneta [22], Minc [18], and Horn and Johnson [12] for more results on non-negative matrices.

The Krein–Rutman theorem

The Perron–Frobenius theorem has several infinite-dimensional extensions. Below we present the Krein–Rutman theorem, which is an infinite-dimensional version of this theorem in some ordered Banach spaces endowed with a solid positive cone, that is, a positive cone with non-empty interior.

To present this result, let us first recall the definition of a compact operator.

Definition 6.5

Let $(X, \| \cdot \|_X)$ be a Banach space. Let $A \in \mathcal{L}(X)$ be a bounded linear operator on the Banach space X. Then A is said to be a *compact operator* if the closure of $A(B)$ is a compact subset of X for every bounded subset B of X.

As mentioned above, there are several extensions of the Perron–Frobenius theorem in Banach spaces. The main assumption of the Krein–Rutman theorem is that the interior of the positive cone X_+ is non-empty. We refer to [1, 3, 14, 26, 21, 23] for various proofs and more results on this subject.

Theorem 6.6 (Krein–Rutman)

Let $(X, \| \cdot \|_X)$ *be a Banach space partially ordered with a positive cone* X_+ *with non-empty interior. Let* $A \in \mathcal{L}(X)$ *be a compact bounded linear operator on* X *such that*

$$Ax \gg 0,\ \forall x > 0 \text{ or equivalently } A(X_+ \setminus \{0\}) \subset \mathring{X}_+.$$

Then the following properties are satisfied

(i) *The spectral radius of* A*, which is defined by* $r(A) = \lim_{n\to\infty} \|A^n\|_{\mathcal{L}(X)}^{1/n}$*, is strictly positive and is a simple eigenvalue of* A *(i.e.* $\dim(N(r(A)I - A))) = 1$ *and* $N((r(A)I - A)) = N\left((r(A)I - A)^2\right)$*).*

(ii) *There exists a* $u \in \mathrm{Int}(X_+)$ *such that* $Au = r(A)u$*.*

(iii) *If* $\lambda \in \mathbb{C}$ *is an eigenvalue of* A *with* $|\lambda| = r(A)$*, then* $\lambda = r(A)$*.*

(iv) *If* $\lambda \in \mathbb{C}$ *is an eigenvalue of* A *such that there exists a* $u \in X_+ \setminus \{0\}$ *with* $Au = \lambda u$*, then* $\lambda = r(A)$*.*

Note that the compactness of the linear operator A can be weakened by working with the essential spectral radius and the measure of non-compactness. More general versions of the above result can also be stated for Banach spaces partially ordered by a cone with an empty interior. We refer the interested readers to the monographs of Schaefer [21], and to those of Meyer-Nieberg [17]. See also [16, 19].

Random linear systems

To quote a few results involving random linear difference equations, we refer to the paper of Cohen and Newman [2] for a nice discussion of examples and counterexamples related to the growth of solutions of a non-autonomous difference equation of the form

$$N(n+1) = A(n)N(n), \forall n \in \mathbb{N}, \text{ and } N(0) = x \in \mathbb{R}^n,$$

where $A(n)$ is an n by n random matrix (i.e. all elements are random variables).

The authors investigate the growth of the solution, that is, the function of x:

$$\lambda(x) = \lim_{n \to +\infty} \frac{\ln(\|N(n)\|)}{n}.$$

For the most advanced readers, we also refer to the work of Lian and Lu [15] on Lyapunov exponents and invariant manifolds for random dynamical systems in a Banach space.

Thank you for listening

- Alikakos, N.D., Fusco, G.: A dynamical systems proof of the Krein–Rutman theorem and an extension of the Perron theorem.
  Proceedings of the Royal Society of Edinburgh Section A: Mathematics **117**(3-4), 209–214 (1991)
- Cohen, J.E., Newman, C.M.: The stability of large random matrices and their products.
  The Annals of Probability pp. 283–310 (1984)
- Deimling, K.: Nonlinear functional analysis.
  Springer-Verlag, Berlin (1985)
- Demetrius, L.: Primitivity conditions for growth matrices.
  Mathematical Biosciences **12**(1-2), 53–58 (1971)
- Diekmann, O., Heesterbeek, J.A.P., Metz, J.A.: On the definition and the computation of the basic reproduction ratio r_0 in models for infectious diseases in heterogeneous populations.
  Journal of mathematical biology **28**(4), 365–382 (1990)

- Van den Driessche, P., Watmough, J.: Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission.
  Mathematical Biosciences **180**(1-2), 29–48 (2002)
- Frobenius, G.F.: Ueber matrizen aus positiven elementen, sitzungsber.
  Preus. Akad. Wiss Berlin pp. 471–476 (1908)
- Frobenius, G.F.: Über matrizen aus positiven elementen, 2?, sitzungsber.
  Königl. Preuss. Akad. Wiss pp. 514–518 (1909)
- Gantmacher, F.R.: The theory of matrices, vol. 2.
  Chelsea Pub Co., New York (1959)
- Greenhalgh, S.: Missile science, population science: the origins of China's one-child policy.
  The China Quarterly **182**, 253–276 (2005)
- Greenhalgh, S.: Just one child: Science and policy in Deng's China.

Univ of California Press (2008)

- Horn, R.A., Johnson, C.R.: Matrix analysis.
  Cambridge University Press (2012)

- Kato, T.: Perturbation theory for linear operators.
  Classics in Mathematics. Springer-Verlag, Berlin (1995).
  Reprint of the 1980 edition

- Krein, M.G., Rutman, M.A.: Linear operators leaving invariant a cone in a Banach space.
  Uspekhi Matematicheskikh Nauk **3**(1), 3–95 (1948)

- Lian, Z., Lu, K.: Lyapunov exponents and invariant manifolds for random dynamical systems in a Banach space.
  American Mathematical Soc. (2010)

- Mallet-Paret, J., Nussbaum, R.D.: Generalizing the Krein–Rutman theorem, measures of noncompactness and the fixed point index.
  Journal of Fixed Point Theory and Applications **7**(1), 103–143 (2010)

Meyer-Nieberg, P.: Banach lattices.
Universitext. Springer-Verlag, Berlin (1991).
DOI 10.1007/978-3-642-76724-1.
URL https://doi.org/10.1007/978-3-642-76724-1

Minc, H.: Nonnegative matrices.
Wiley-Interscience Series in Discrete Mathematics and Optimization.
John Wiley & Sons, Inc., New York (1988)

Nussbaum, R.D.: Eigenvectors of nonlinear positive operators and the linear krein–rutman theorem.
In: Fixed point theory, pp. 309–330. Springer (1981)

Perron, O.: Zur Theorie der Matrices.
Math. Ann. **64**(2), 248–263 (1907).
DOI 10.1007/BF01449896.
URL https://doi.org/10.1007/BF01449896

Schaefer, H.H.: Banach lattices.

In: Banach Lattices and Positive Operators, pp. 46–153. Springer (1974)

Seneta, E.: Non-negative matrices and Markov chains. Springer Science & Business Media (2006)

Smoller, J.: Shock waves and reaction-diffusion equations, volume 258 of. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] (1983)

Song, J.: Some developments in mathematical demography and their application to the People's Republic of China. Theoretical Population Biology **22**(3), 382–391 (1982)

Song, J., Kong, D., Yu, J.: Population system control. Mathematical and Computer Modelling **11**, 11–16 (1988)

Takác, P.: A short elementary proof of the Krein–Rutman theorem. Houston Journal of Mathematics **20**(1), 93–98 (1994)

Thieme, H.R.: Spectral bound and reproduction number for infinite-dimensional population structure and time heterogeneity. SIAM Journal on Applied Mathematics **70**(1), 188–211 (2009)