# On the Numerical Characterization of the Reachability Cone for an Essentially Nonnegative Matrix


Dimitrios Noutsos
Department of Mathematics
University of Ioannina
GR - 451 10
Ioannina, Greece
(dnoutsos@uoi.gr)

Michael J. Tsatsomeros
Mathematics Department
Washington State University
Pullman, WA 99164-3113
(tsat@wsu.edu)


October 22, 2008


**Abstract**

Given an $n \times n$ real matrix $A$ with nonnegative off-diagonal entries, the solution to $\dot{x}(t) = Ax(t)$, $x_0 = x(0)$, $t \geq 0$ is $x(t) = e^{tA}x_0$. The problem of identifying the initial points $x_0$ for which $x(t)$ becomes and remains entrywise nonnegative is considered. It is known that such $x_0$ are exactly those vectors for which the iterates $x^{(k)} = (I + hA)^k x_0$ become and remain nonnegative, where $h$ is a positive, not necessarily small parameter that depends on the diagonal entries of $A$. In this paper, this characterization of initial points is extended to a numerical test when $A$ is irreducible: If $x^{(k)}$ becomes and remains positive, then so does $x(t)$; if $x(t)$ fails to become and remain positive, then either $x^{(k)}$ becomes and remains negative or it always has a negative and a positive entry. Due to round-off errors, the latter case manifests itself numerically by $x^{(k)}$ converging with a relatively small convergence ratio to a positive or a negative vector. An algorithm implementing this test is provided, along with its numerical analysis and examples. The reducible case is also discussed and a similar test is described.



**Key words:** Essentially nonnegative matrix; exponentially nonnegative matrix; reachability cone; Perron-Frobenius; power method.

**AMS subject classifications:** 15A48, 34A30, 34C99, 65F10.

# 1 Introduction

In this paper we study the dynamics associated with the linear differential system

$$\dot{x}(t) = Ax(t), \quad A \in \mathbb{R}^{n\times n}, \quad x(0) = x_0 \in \mathbb{R}^n, \quad t \geq 0, \tag{1.1}$$

where the coefficient matrix $A$ is essentially nonnegative, i.e., it has nonnegative off-diagonal entries. Such systems arise frequently in applications, for example in engineering and mathematical biology among others; see [1, 2, 7]. The solution to the system (1.1) is given by

$$x(t) = e^{tA}x_0, \quad \forall t \geq 0.$$

We shall refer to the set $\{x(t) = e^{tA}x_0 \mid t \in [0, \infty)\}$ as the *continuous trajectory emanating from* $x_0$ and say that $x_0$ gives rise to this trajectory. As the main concern of this presentation, we pose the following 'hit and hold' question: *When does the trajectory emanating from a given initial point $x_0$ become (entrywise) nonnegative and remain such for all time thereafter?* In other words, we aim to identify all initial points $x_0$ for which there exists a finite time $t_0 \geq 0$ such that $e^{tA}x_0 \in \mathbb{R}^n_+$ for all $t \geq t_0$, where $\mathbb{R}^n_+$ denotes the set of all entrywise nonnegative vectors in $\mathbb{R}^n$. The set of all such points is known as the reachability cone of $\mathbb{R}^n_+$, denoted by $X_A(\mathbb{R}^n_+)$.

Our efforts herein are toward a numerical characterization of the members of $X_A(\mathbb{R}^n_+)$. We shall build our work on results previously established in [6]. It is known that $x_0 \in X_A(\mathbb{R}^n_+)$ if and only if the iterates $x^{(k)} = (I + hA)^k x_0$ become and remain nonnegative, where $h$ is some positive, not necessarily small parameter that depends on the diagonal entries of $A$. When $A$ is irreducible, we develop a comprehensive numerical test as follows. If $x^{(k)}$ becomes positive, then so does $x(t)$; if $x(t)$ does not become and remain positive, then two possibilities exist: either $x^{(k)}$ becomes and remains negative or it always has a negative and a positive entry. Due to round-off errors, the latter case manifests itself numerically by $x^{(k)}$ converging with a relatively small convergence ratio to a positive or negative vector. An algorithm implementing this test is provided, along with its theoretical basis, numerical analysis and illustrative examples. The reducible case is also discussed and a possible similar test is described.

Section 2 contains definitions and notation used throughout the paper. In Section 3 we describe the continuous and discrete reachability cones of $\mathbb{R}^n_+$ associated with an essentially nonnegative matrix, as well as review the necessary material from [6]. In Section 4 we study the relation between continuous and discrete trajectories in the irreducible case. In Section 5 we provide and analyze numerically an algorithm to decide membership in $X_A(\mathbb{R}^n_+)$ when $A$ is irreducible. Section 6 contains numerical examples. Finally, in Section 7 we discuss the general (possibly reducible) case and describe what is entailed in adapting our algorithm to this case.

## 2 Definitions and notation

Given a vector $x \in \mathbb{R}^n$, $(x)_i$ denotes the $i$-th entry of $x$. The *nonnegative orthant*, $\mathbb{R}^n_+$ is the set of all (entrywise) nonnegative vectors in $\mathbb{R}^n$. The topological interior of $\mathbb{R}^n_+$ is denoted by $\mathrm{int}\mathbb{R}^n_+$. We use the notations $x \geq 0$ ($x > 0$) and $x \in \mathbb{R}^n_+$ ($x \in \mathrm{int}\mathbb{R}^n_+$) interchangeably.

Given an $n \times n$ matrix $Y$, the spectrum of $Y$ is denoted by $\sigma(Y)$ and its spectral radius by $\rho(Y)$. The *spectral abscissa* of $Y$ is defined and denoted by $\lambda(Y) = \max\{\mathrm{Re}(\lambda) \mid \lambda \in \sigma(Y)\}$. An eigenvalue $\lambda$ of $Y$ is said to be *dominant* if $|\lambda| = \rho(Y)$. By $\mathrm{index}_\lambda(Y)$ we denote the degree of $\lambda$ as a root of the minimal polynomial of $Y$.

**Definition 2.1** An $n \times n$ matrix $Y = [y_{ij}]$ is called:

- *nonnegative* (*positive*), denoted by $Y \geq 0$ ($> 0$), if $y_{ij} \geq 0$ ($> 0$) for all $i$ and $j$;
- *essentially nonnegative* (*positive*), denoted by $Y \overset{\mathrm{e}}{\geq} 0$ ($Y \overset{\mathrm{e}}{>} 0$), if $y_{ij} \geq 0$ ($> 0$) for all $i \neq j$;
- *reducible* if there exists a permutation matrix $P$ such that
$$PYP^T = \begin{pmatrix} Y_{11} & 0 \\ Y_{21} & Y_{22} \end{pmatrix},$$
where $Y_{11}$ and $Y_{22}$ are square, non-vacuous matrices. Otherwise, $Y$ is called *irreducible*.
- *primitive* if $Y \geq 0$ and there exists positive integer $m$ such that $Y^m > 0$.

In the following theorem, we summarize for the sake of reference the basic premises of the Perron-Frobenius Theorem; see [2, Chapter 2,Theorems (1.1) and (1.3)] or [3, Chapter 8].

**Theorem 2.2** *Let $B \in \mathbb{R}^{n\times n}$ be a nonnegative matrix. Then the following hold:*

(i) *$\rho(B) \in \sigma(B)$ and there is a nonnegative eigenvector corresponding to $\rho(B)$.*

(ii) *If $B$ is irreducible, then $\rho(B)$ is a simple eigenvalue having a positive eigenvector $w$. In addition, every nonnegative eigenvector of $B$ is a multiple of $w$.*

(iii) *If $B$ is irreducible, then $(I + B)^k > 0$ for all $k \geq n - 1$.*

**Definition 2.3** Let $Y \in \mathbb{R}^{n\times n}$ and $\mu \in \sigma(Y)$. The *generalized eigenspace* of $Y$ *corresponding to* $\mu$ is the $Y$-invariant subspace

$$\mathcal{N}_\mu(Y) = \mathrm{Nul}\,(Y - \mu I)^{\mathrm{index}_\mu(Y)}.$$

Let $B \geq 0$ be irreducible. By Theorem 2.2, $\rho(B) \in \sigma(B)$ is simple and $N_{\rho(B)}(B)$ is spanned by an eigenvector $w \in \mathrm{int}\mathbb{R}^n_+$. We refer to $N_{\rho(B)}(B)$ as the *Perron eigenspace* of $B$ and to $w$ as the *Perron vector* of $B$. We also denote

$$\mathcal{L}_B \;=\; \bigoplus_{\mu \neq \rho(B)} \mathcal{N}_\mu(B),$$

where the direct sum is taken over all distinct eigenvalues $\mu$ of $B$ with $\mu \neq \rho(B)$. Recall that

$$\mathbb{C}^n \;=\; N_{\rho(B)}(B) \bigoplus \mathcal{L}_B.$$

The same terminology and similar notation are used for an irreducible $A \overset{\mathrm{e}}{\geq} 0$ and its eigenvalue $\lambda(A)$.

**Definition 2.4** For an $n \times n$ matrix $A = [a_{ij}] \overset{\mathrm{e}}{\geq} 0$, we define the quantity

$$h(A) = \sup\{\, h \mid \min_{1 \leq i \leq n} (1 + h a_{ii}) > 0\}.$$

Notice that $h(A) = \sup\{h \mid (I + hA) \geq 0\} \geq 0$ and that $h(A) = \infty$ when $A \geq 0$.

**Definition 2.5** An $n \times n$ matrix $A$ is called *exponentially nonnegative* if

$$e^{tA} = \sum_{i=0}^{\infty} \frac{t^k A^k}{k!} \geq 0, \quad \forall t \geq 0.$$

If $e^{tA} > 0$ for all $t > 0$, $A$ is called *exponentially positive.*

**Lemma 2.6** [2, Chapter 6, Theorem (3.12)] *An $n \times n$ matrix $A$ is exponentially nonnegative if and only if $A \overset{\mathrm{e}}{\geq} 0$. Moreover, if $A$ is irreducible, then $A$ is exponentially positive if and only if $A \overset{\mathrm{e}}{>} 0$.*

**Definition 2.7** Consider a set $\Gamma \subset \mathbb{R}^n$ and $Y \in \mathbb{R}^{n \times n}$. $\Gamma$ is called

- $Y$*-invariant* (or invariant under $Y$) if $Y\Gamma \subseteq \Gamma$.
- *positively invariant* with respect to $Y \in \mathbb{R}^{n \times n}$ if $e^{tY}\,\Gamma \subseteq \Gamma$, $\forall t \geq 0$.

Positive invariance of $\Gamma$ has the implication that once a trajectory emanating from $x_0$ reaches $\Gamma$ in some finite time, it remains in $\Gamma$ for all finite time thereafter.

**Definition 2.8** Let $\Gamma$ be a positively invariant set with respect to $A \in \mathbb{R}^{n\times n}$. The set of all initial points of trajectories which reach and remain $\Gamma$ is referred to as the *reachability set* of $\Gamma$ under $A$ and is denoted by $X_A(\Gamma)$; that is,

$$X_A(\Gamma) \;=\; \{x_0 \in \mathbb{R}^n \mid (\exists t_0 = t(x_0) \geq 0)\,(\forall t \geq t_0)\,[e^{tA}x_0 \in \Gamma]\} \;=\; \bigcup_{t\geq 0} e^{-tA}\Gamma. \tag{2.1}$$

Let $\Gamma \subseteq \mathbb{R}^n$ be a proper cone, i.e., a closed, pointed, solid convex set; see [2, Chapter 1] and [8]. For example, $\mathbb{R}^n_+$ is a proper cone in $\mathbb{R}^n$. Then, as shown in [5], $X_A(\Gamma)$ is a convex cone that contains $\Gamma$, but is not necessarily closed or pointed.

# 3 The reachability cones

Consider an $n \times n$ essentially nonnegative matrix $A$ and the (continuous) *reachability cone* (of $\mathbb{R}^n_+$ under $A$),

$$X_A(\mathbb{R}^n_+) \;=\; \{x_0 \in \mathbb{R}^n \mid (\exists t_0 = t(x_0) \geq 0)\,(\forall t \geq t_0)\,[e^{tA}x_0 \in \mathbb{R}^n_+]\}.$$

In what follows, we describe results from [6] on which we shall base our theoretical analysis and numerical characterization of $X_A(\mathbb{R}^n_+)$.

First, we refer to the the sequence $\{x^{(k)}\}$ generated from $x_0$ by the Cauchy-Euler finite differences scheme

$$x^{(k)} = (I + hA)^k x_0, \;\; k = 0, 1, \ldots$$

as *the discrete trajectory* (associated with the time-step $h$) *emanating from* $x_0 = x^{(0)}$.

Second, given an essentially nonnegative matrix $A$ and any $h \in (0, h(A))$, we denote by $X_{A,h}(\mathbb{R}^n_+)$ the set of all initial states $x_0 \in \mathbb{R}^n$ which give rise to discrete trajectories $\{x^{(k)}\}$ that become (and remain, due to nonnegativity of $I + hA$) nonnegative; that is,

$$X_{A,h}(\mathbb{R}^n_+) = \{x_0 \in \mathbb{R}^n \mid (\exists k_0 = k_0(x_0) \geq 0)(\forall k \geq k_0)[(I + hA)^k x_0 \in \mathbb{R}^n_+]\}.$$

We refer to $X_{A,h}(\mathbb{R}^n)$ as the *discrete reachability cone* (of $\mathbb{R}^n_+$ under $A$ with respect to $h$).

The continuous and discrete reachability sets of $\mathrm{int}\mathbb{R}^n_+$ under $A$ are defined analogously by requiring trajectories to become and remain positive; they are denoted by $X_A(\mathrm{int}\mathbb{R}^n_+)$ and $X_{A,h}(\mathrm{int}\mathbb{R}^n_+)$, respectively.

The geometric and algebraic properties of (continuous and discrete) reachability cones are studied extensively in [4, 5, 6]. Next is a summary of results from [6], expressed here for both the reachability of $\mathbb{R}^n_+$ and of $-\mathbb{R}^n_+$.

**Theorem 3.1** [6] *Let $A$ be an $n \times n$ essentially nonnegative matrix and let $h \in (0, h(A))$ such that $(I + hA)$ is invertible. Then*

$$X_A(\mathbb{R}^n_+) = X_{A,h}(\mathbb{R}^n_+) \quad \textit{and} \quad X_A(-\mathbb{R}^n_+) = X_{A,h}(-\mathbb{R}^n_+). \tag{3.1}$$

*If, in addition, $A$ is irreducible, then*

$$X_A(\mathbb{R}^n_+) \setminus \{0\} = X_A(\mathrm{int}\mathbb{R}^n_+) = X_{A,h}(\mathrm{int}\mathbb{R}^n_+) = X_{A,h}(\mathbb{R}^n_+) \setminus \{0\} \tag{3.2}$$

*and*

$$X_A(-\mathbb{R}^n_+) \setminus \{0\} = X_A(-\mathrm{int}\mathbb{R}^n_+) = X_{A,h}(-\mathrm{int}\mathbb{R}^n_+) = X_{A,h}(-\mathbb{R}^n_+) \setminus \{0\}.$$

**Proof.** The first equation in (3.1) is [6, Theorem 3.3]. The second equation in (3.1) follows simply by observing that $X_A(-\mathbb{R}^n_+) = -X_A(\mathbb{R}^n_+)$ and similarly for the discrete reachability cones.

If $A$ is irreducible, it follows by Theorem 2.2 that $(I + hA)^k > 0$ for all $k \geq n - 1$. Also by Lemma 2.6, $e^{tA} > 0$ for all $t > 0$. Consequently,

$$\begin{aligned}
X_A(\mathbb{R}^n_+) \setminus \{0\} &= X_A(\mathrm{int}\mathbb{R}^n_+) \\
&= \mathrm{int}X_A(\mathbb{R}^n_+) \quad \text{(by [6, Lemma 3.2])} \\
&= \mathrm{int}X_{A,h}(\mathbb{R}^n_+) \quad \text{(by (3.1))} \\
&= X_{A,h}(\mathrm{int}\mathbb{R}^n_+) \quad \text{(by [6, Lemma 3.2])} \\
&= X_{A,h}(\mathbb{R}^n_+) \setminus \{0\},
\end{aligned}$$

proving (3.2). The last equation follows analogously to the last equation in (3.1). ∎

Theorem 3.1 suggests a numerical test to determine whether a given initial point $x_0$ belongs to $X_A(\mathbb{R}^n_+)$ or not:

**1.** choose a positive $h < h(A)$ such that the iteration matrix $I + hA$ is invertible;

**2.** check whether for some nonnegative integer $k$, $x^{(k)} = (I + hA)^k x_0$ is nonnegative (in which case $x_0 \in X_A(\mathbb{R}^n_+)$) **or** decide that $x^{(k)}$ will never become nonnegative (in which case $x_0 \notin X_A(\mathbb{R}^n_+)$).

Some questions arise immediately regarding this approach: How can one decide whether $x^{(k)}$ will never become nonnegative? What are the numerical effects of the iterative generation of $x^{(k)}$ in making such a decision? We begin the task of answering these questions next.

## 4 Main theoretical results in the irreducible case

Here we provide further analysis of the relation between continuous and discrete trajectories that will lead to an algorithmic characterization of $X_A(\mathbb{R}^n_+)$ when $A$ is irreducible.

**Lemma 4.1** *Let $B$ be an $n \times n$ irreducible nonnegative matrix. Then $\mathcal{L}_B = R(\rho(B)I - B)$, where $R(\cdot)$ denotes the range.*

**Proof.**

Let $\{w_1, w_2, \ldots, w_n\}$ be a basis for $\mathbb{C}^n$, where $w_1$ is the Perron vector of $B$ and where $\mathrm{Span}\{w_2, w_3, \ldots, w_n\} = \mathcal{L}_B$. Let $x \in R(\rho(B)I - B)$ so that

$$x = (\rho(B)I - B)y, \quad \text{where} \quad y = \sum_{j=1}^{n} c_j\, w_j, \quad c_j \in \mathbb{C} \;\; (j = 1, 2, \ldots, n).$$

That is, as $\mathcal{L}_B$ is a $B$-invariant subspace, there exist $q_j \in \mathbb{C}$ ($j = 2, 3, \ldots, n$) such that

$$x = \sum_{j=2}^{n} q_j\, w_j \in \mathcal{L}_B.$$

Thus $R(\rho(B)I - B) \subseteq \mathcal{L}_B$. As $\rho(B)$ is a simple eigenvalue, both subspaces have dimension $n - 1$ and so $\mathcal{L}_B = R(\rho(B)I - B)$. ∎

We continue with a result on primitive matrices of general interest. Recall that for an irreducible nonnegative matrix, primitivity is equivalent to the spectral radius being the sole dominant eigenvalue; see [2, Chapter 2, Theorem (1.7)].

**Proposition 4.2** *Let $B$ be an $n \times n$ irreducible nonnegative matrix such that $\rho(B) > |\mu|$ for all eigenvalues $\mu$ of $B$ with $\mu \neq \rho(B)$. Then $\mathcal{L}_B \cap \mathbb{R}^n_+ = \{0\}$.*

**Proof.**

Let us first show that $\mathcal{L}_B$ does not contain any positive vectors. By way of contradiction, suppose that $0 < x \in \mathcal{L}_B$. By Lemma 4.1, there exists $y \in \mathbb{R}^n$ such that $x = (\rho(B)I - B)y > 0$. Then, for all sufficiently small $\epsilon > 0$,

$$((\rho(B) + \epsilon)I - B)\, y = x + \epsilon y > 0.$$

Since $C = (\rho(B) + \epsilon)I - B$ is a nonsingular M-matrix, we have that $C^{-1} > 0$; see [2, Chapter 6, Theorem (2.3)]. Thus $y = C^{-1}(x + \epsilon y) > 0$ and as $x > 0$ we have that

$$\rho(B)y > By, \quad y > 0.$$

The latter inequalities imply that $\rho(B) > \rho(B)$; see [3, Corollary 8.1.29]. This is a contradiction showing that $\mathcal{L}_B \cap \mathrm{int}\mathbb{R}^n_+ = \emptyset$.

To prove the theorem's assertion, let now $0 \neq x \in \mathcal{L}_B \cap \mathbb{R}^n_+$. As $B$ is primitive and $\mathcal{L}_B$ is $B$-invariant, $B^{n-1}x \in \mathcal{L}_B \cap \mathrm{int}\mathbb{R}^n_+ = \emptyset$; a contradiction proving that $\mathcal{L}_B \cap \mathbb{R}^n_+ = \{0\}$. ∎

In essence, the next lemma proves formally a simple fact (stated here in the context of essentially nonnegative matrices): When a matrix $Y$ has a dominant eigenvalue $\mu > 0$ and when $\alpha > 0$, then (up to algebraic multiplicity) the only dominant eigenvalue of $Y + \alpha I$ is $\mu + \alpha$.

**Lemma 4.3** *Let $A \in \mathbb{R}^{n\times n}$ be an essentially nonnegative matrix. Then for every $h \in (0, h(A))$ we have that $1 + h\lambda(A) > |1 + h\mu|$ for all $\mu \in \sigma(A)$ with $\mu \neq \lambda(A)$.*

**Proof.** Let $\alpha(A)$:=$\min\{\alpha \geq 0 \mid A + \alpha I \geq 0\}$ and notice that as $\alpha(A) = h(A)^{-1}$, in order to prove this lemma's assertion, it is sufficient to show that $\lambda(A) + \alpha > |\mu + \alpha|$ for all $\mu \in \sigma(A) \setminus \{\lambda(A)\}$ and all $\alpha > \alpha(A)$.

Clearly, $\rho(A + \alpha(A)I) = \lambda(A) + \alpha(A)$ and so $\lambda(A) + \alpha(A) \geq |\mu + \alpha(A)|$ for all $\mu \in \sigma(A) \setminus \{\lambda(A)\}$; we wish to show that this inequality is strict for all $\alpha > \alpha(A)$.

If $\lambda(A) + \alpha(A) > |\mu + \alpha(A)|$ for all $\mu \in \sigma(A) \setminus \{\lambda(A)\}$, we are done. Otherwise, there exists $\omega \in \sigma(A) \setminus \{\lambda(A)\}$ such that

$$\lambda(A) \ + \ \alpha(A) = |\omega \ + \ \alpha(A)|.$$

Set $\tau = \lambda(A) + \alpha(A)$, $\xi = \omega + \alpha(A)$, consider $\varepsilon > 0$ arbitrary, and compare $\tau + \varepsilon$ to $|\xi + \varepsilon|$. If $\omega \in \mathbb{R}$, since $\omega \neq \lambda(A)$, we have $\tau = -\xi$. Then

$$\tau + \varepsilon = -\xi + \varepsilon > -\xi - \varepsilon = |\xi + \epsilon|.$$

When $\omega = a + ib$ with $b \neq 0$, we have

$$|\xi \ + \ \varepsilon|^2 = (a \ + \ \alpha(A) \ + \ \varepsilon)^2 \ + \ b^2 = (a \ + \ \alpha(A))^2 + 2(a \ + \ \alpha(A))\varepsilon \ + \ \varepsilon^2 \ + \ b^2. \quad (4.1)$$

Since $\tau = |\xi|$ by assumption, $\tau^2 = (a + \alpha(A))^2 + b^2$ and as a consequence of (4.1),

$$|\xi \ + \ \varepsilon|^2 = \tau^2 \ + \ \varepsilon^2 \ + \ 2(a \ + \ \alpha(A))\varepsilon.$$

However, $(\tau \ + \ \varepsilon)^2 = \tau^2 \ + \ \varepsilon^2 \ + \ 2\tau\varepsilon$ and since $\tau = |\xi| > \mathrm{Re}(\xi) = a + \alpha(A)$, it follows that

$$\lambda(A) \ + \ \alpha = \lambda(A) \ + \ (\alpha(A) + \varepsilon) = \tau \ + \ \varepsilon \ > \ |\xi \ + \ \varepsilon| = |\omega \ + \ (\alpha(A) + \varepsilon)| = |\omega \ + \ \alpha|.$$

Thus $\lambda(A) + \alpha > |\mu + \alpha|$ for all $\mu \in \sigma(A) \setminus \{\lambda(A)\}$ and $\alpha > \alpha(A)$. ∎

**Remark 4.4** Let $A \stackrel{\text{e}}{\geq} 0$. Then the following are useful observations.

1. In view of Lemma 4.3 and Theorem 2.2, if $A$ is irreducible, then for every shift $h \in (0, h(A))$, the spectral radius of $I+hA$ is a simple and the sole dominant eigenvalue of $I + hA$. That is, $I + hA$ is a primitive matrix.

2. As $\sigma(I + hA) = 1 + h\sigma(A)$, it is true that for all but a finite number of such shifts $h \in (0, h(A))$, $I + hA$ is an invertible matrix.

**Theorem 4.5** *Let $A$ be an $n \times n$ irreducible essentially nonnegative matrix and consider an $h \in (0, h(A))$ such that $B = (I + hA)$ is invertible. Let $\Gamma = \mathbb{R}^n_+ \cup (-\mathbb{R}^n_+)$. Then*

$$X_A(\text{int}\Gamma) \cup \{0\} \;=\; X_A(\Gamma) \;=\; (\mathbb{R}^n \setminus \mathcal{L}_B) \cup \{0\}.$$

**Proof.**

Let $A$ and $h$ as prescribed in the statement of the theorem and let $w$ be the Perron vector of $B = I + hA$, and thus of $A$. Since $B \geq 0$, $\mathbb{R}^n_+$ and $-\mathbb{R}^n_+$ are invariant under $B$. Also, by Lemma 2.6, $\mathbb{R}^n_+$ and $-\mathbb{R}^n_+$ are positively invariant under $A$. That is,

$$\begin{aligned}
X_A(\Gamma) &= X_A\left(\mathbb{R}^n_+ \cup (-\mathbb{R}^n_+)\right) \\
&= X_A(\mathbb{R}^n_+) \cup X_A(-\mathbb{R}^n_+) \quad \text{(since } \mathbb{R}^n_+ \text{ and } -\mathbb{R}^n_+ \text{ are positively invariant)} \\
&= X_{A,h}(\mathbb{R}^n_+) \cup X_{A,h}(-\mathbb{R}^n_+) \quad \text{(by Theorem 3.1)} \\
&= X_{A,h}(\text{int}\mathbb{R}^n_+) \cup X_{A,h}(-\text{int}\mathbb{R}^n_+) \cup \{0\} \quad \text{(since } B = I + hA \text{ is primitive)} \\
&= X_A(\text{int}\mathbb{R}^n_+) \cup X_A(-\text{int}\mathbb{R}^n_+) \cup \{0\} \quad \text{(by Theorem 3.1)} \\
&= X_A(\text{int}\Gamma) \cup \{0\}.
\end{aligned}$$

Next, recall that $\mathbb{C}^n = \bigoplus_{\mu \in \sigma(B)} \mathcal{N}_\mu(B)$. Therefore, for any nonzero initial vector $x_0 \in \mathbb{R}^n \setminus \mathcal{L}_B$, by the premises of the power method [10], and by Remark 4.4 part 1, $B^k x_0$ converges as $k \to \infty$ to $cw$, where $c$ is a nonzero real number. As a consequence, the discrete (and thus by Theorem 3.1 the continuous) trajectories emanating from $x_0$ will necessarily either enter and remain in $\mathbb{R}^n_+$ (if $c > 0$) or enter and remain in $-\mathbb{R}^n_+$ (if $c < 0$). That is,

$$(\mathbb{R}^n \setminus \mathcal{L}_B) \cup \{0\} \;\subseteq\; X_A(\Gamma).$$

By Lemma 4.3, $\rho(B)$ is the sole dominant eigenvalue of $B$. Thus, by Proposition 4.2, $\mathcal{L}_B \cap \text{int}\mathbb{R}^n_+ = \emptyset$. It follows that

$$X_A(\Gamma) \;\subseteq\; (\mathbb{R}^n \setminus \mathcal{L}_B) \cup \{0\},$$

completing the proof of the theorem. ∎

## 5 Algorithmic characterization of $X_A(\mathbb{R}^n_+)$ for irreducible $A$

Our goal in this section is to develop and analyze an iterative algorithm for the detection of the members of $X_A(\mathbb{R}^n_+)$ based on the following interpretation of Theorem 4.5.

**Theorem 5.1** *Let $A$ be an $n \times n$ irreducible essentially nonnegative matrix and consider an $h \in (0, h(A))$ such that $B = (I + hA)$ is invertible. For any $x_0 \in \mathbb{R}^n$, consider the vectors $x^{(k)} = B^k x_0$ $(k = 0, 1, \ldots)$. Then exactly one of the following three alternatives occurs:*

(1) *$x_0 \in X_A(\mathbb{R}^n_+)$ and there exists integer $k_0 \geq 0$ such that $x^{(k)} > 0$ for all $k \geq k_0$.*

(2) *$x_0 \notin X_A(\mathbb{R}^n_+) \cup \mathcal{L}_B$ and there exists integer $k_0 \geq 0$ such that $x^{(k)} < 0$ for all $k \geq k_0$.*

(3) *$x \in \mathcal{L}_B$ and for all integers $k \geq 0$, $x^{(k)}$ has a positively and a negatively signed entry.*

Throughout this section, $A$ is an $n \times n$ irreducible nonnegative matrix and $B = I + hA$, where $h \in (0, h(A))$ is chosen so that $B$ is invertible. In view of Remark 4.4 part 1, $B$ is an irreducible nonnegative matrix whose spectral radius $\rho(B)$ is simple and the sole dominant eigenvalues of $B$. As a consequence, we can consider the eigenvalues of $B$ to be ordered as

$$\mu_1 > |\mu_2| \geq |\mu_3| \geq \ldots \geq |\mu_n|,$$

where $\mu_1 = \rho(B) = 1 + h\lambda(A)$.

Let $w_1, w_2, \ldots, w_n$ be a Jordan basis of $\mathbb{C}^n$ consisting of generalized eigenvectors of $B$ (and thus of $A$). In this basis, we take $w_1 > 0$ to be the Perron vector of $B$ corresponding to the simple eigenvalue $\mu_1$. It follows that

$$\mathrm{Span}\{w_2, w_3 \ldots, w_n\} \;=\; \mathcal{L}_A = \bigoplus_{\mu \neq \rho(B)} \mathcal{N}_\mu(B).$$

Let us now consider an initial vector

$$x_0 = c_1 w_1 + c_2 w_2 + \cdots + c_n w_n \in \mathbb{R}^n \quad \text{where} \quad c_j \in \mathbb{C},\; j = 1, 2, \ldots, n, \tag{5.1}$$

and the sequence of iterates

$$x^{(k)} = B^k x_0 = c_1 B^k w_1 + c_2 B^k w_2 + \cdots + c_n B^k w_n, \quad k = 0, 1, \ldots. \tag{5.2}$$

If $c_1 \neq 0$ in (5.1), by the premises of the Power Method (see, e.g., [10]), the sequence $\{x^{(k)}\}$ converges to the subspace spanned by $w_1$ with convergence ratio $\frac{\mu_2}{\mu_1}$. If $\{x^{(k)}\}$ becomes a positive vector, we conclude that $x_0 \in X_A(\mathbb{R}^n_+)$. If $\{x^{(k)}\}$ becomes a negative vector, we conclude that $x_0 \notin X_A(\mathbb{R}^n_+)$.

If $c_1 = 0$ in (5.1), then, theoretically, $x^{(k)} \in \mathcal{L}_B$ (for all $k = 0, 1, \ldots$). In practice, however, a direction along $w_1$ will be introduced in $x^{(k)}$ due to round-off errors. That is, for $k \geq 1$, $x^{(k)} \in \mathrm{Span}\{w_1, w_2, \ldots, w_k\}$ with the coefficient of $w_1$ being small in magnitude, typically of the order of the machine tolerance, *tol*. In other words, $\{x^{(k)}\}$ will still converge to either a positive or a negative multiple of $w_1$.

As a consequence of the above analysis, in implementing a numerical algorithm to decide whether $x_0$ belongs to $X_A(\mathbb{R}^n_+)$ or not, we need to be able to distinguish whether convergence to a positive vector is due to round-off error or not. More specifically, if $\{x^{(k)}\}$ converges to a positive vector, we need to decide whether $x_0 \in X_A(\mathbb{R}^n_+)$ or $x_0 \in \mathcal{L}_B$. We develop a method to do so next.

Suppose that in (5.2), $x^{(k)} = B^k x_0$ becomes positive or negative at $k = k_0 \geq 0$ for the first time. If $x^{(k_0)} < 0$, then clearly $x_0 \notin X_A(\mathbb{R}^n_+)$. Suppose $x^{(k_0)} > 0$ and thus $x^{(k)} > 0$ for all $k \geq k_0$. Consider then

$$y_0 := x^{(k_0)} = f_1 w_1 + f_2 w_2 + \cdots + f_n w_n, f_1 \in \mathbb{R} \setminus \{0\}, f_j \in \mathbb{C},\ j = 2, 3 \ldots, n, \tag{5.3}$$

as well as the iteration

$$y^{(k)} = B^k y_0, \quad k = 0, 1, \ldots \tag{5.4}$$

**Case 1:** $\mathrm{index}_{\mu_2}(B) = 1$

In this case, letting $m$ be the algebraic multiplicity of $\mu_2$, we have

$$B^k y_0 = f_1 \mu_1^k w_1 + \mu_2^k \sum_{j=2}^{m+1} f_j w_j + \mathcal{O}(\mu_{m+2}^k) w, \tag{5.5}$$

where $w$ belongs to the generalized eigenspace of $\mu_{m+2}$. Since $f_1 \neq 0$, consider (for $k = 1, 2, \ldots$) the quantity

$$\begin{aligned} \widehat{\mu}_1^{(k)} \quad &:= \quad \frac{\left(B^k y_0\right)_i}{(B^{k-1} y_0)_i} = \frac{f_1 \mu_1^k (w_1)_i + \left(\sum_{j=2}^{m+1} f_j w_j\right)_i \mu_2^k + \mathcal{O}(\mu_{m+2}^k)}{f_1 \mu_1^{k-1} (w_1)_i + \left(\sum_{j=2}^{m+1} f_j w_j\right)_i \mu_2^{k-1} + \mathcal{O}(\mu_{m+2}^{k-1})} \\ &= \quad \mu_1 + \frac{\left(\sum_{j=2}^{m+1} f_j w_j\right)_i (\mu_2 - \mu_1)}{f_1 (w_1)_i} \left(\frac{\mu_2}{\mu_1}\right)^{k-1} + \mathcal{O}\left(\left(\frac{\mu_{m+2}}{\mu_1}\right)^{k-1}\right), \end{aligned} \tag{5.6}$$

where $(y_0)_i = \max_j (y_0)_j$. That is, $\widehat{\mu}_1^{(k)}$ provides an estimate to $\mu_1$ with the error of the estimation being $\mathcal{O}\left(\left(\frac{\mu_2}{\mu_1}\right)^{k-1}\right)$.

Taking into account relation (5.2), since $w_1 > 0$ and since $x^{(k_0)} > 0$ (or $< 0$), while $x^{(k_0-1)}$ is neither positive nor negative, we have that there exists index $\ell$ such that

$$\left|c_1 \mu_1^{k_0-1}\,(w_1)_\ell\right| \;\le\; \left|\sum_{j=2}^{m+1} c_j (w_j)_\ell\, \mu_2^{k_0-1}\right|$$

and

$$\left|c_1 \mu_1^{k_0}\,(w_1)_\ell\right| \;>\; \left|\sum_{j=2}^{m+1} c_j (w_j)_\ell\, \mu_2^{k_0}\right|.$$

Thus, for

$$\alpha_j = \frac{(w_j)_\ell}{(w_1)_\ell}, \quad j = 2, 3, \ldots, m+1,$$

we have

$$\left|\frac{\mu_2}{\mu_1}\right|^{k_0} \;<\; \left|\frac{c_1}{\sum_{j=2}^{m+1} \alpha_j c_j}\right| \;\le\; \left|\frac{\mu_2}{\mu_1}\right|^{k_0-1}.$$

That is, $\left(\frac{\mu_2}{\mu_1}\right)^{k_0}$ estimates in magnitude the quantity $\dfrac{c_1}{\sum_{j=2}^{m+1} \alpha_j c_j}$ and thus can serve as a measure of whether $c_1$ is a result of round-off error or not. If it is near the machine tolerance *tol*, then we conclude that $x_0 \in \mathcal{L}_A$; if it is far from *tol*, then we conclude that $x_0 \in X_A(\mathbb{R}^n_+)$ (or $x_0 \notin X_A(\mathbb{R}^n_+) \cup \mathcal{L}_B$).

In order to estimate the ratio $\frac{\mu_2}{\mu_1}$, we consider the difference between two consecutive terms of (5.6):

$$\begin{aligned}
\widehat{\mu}_1^{(k)} - \widehat{\mu}_1^{(k-1)} &= \frac{\left(\sum_{j=2}^{m+1} f_j w_j\right)_i (\mu_2 - \mu_1)}{f_1 (w_1)_i} \left(\frac{\mu_2}{\mu_1}\right)^{k-1} - \frac{\left(\sum_{j=2}^{m+1} f_j w_j\right)_i (\mu_2 - \mu_1)}{f_1 (w_1)_i} \left(\frac{\mu_2}{\mu_1}\right)^{k-2} \\
&\quad + \mathcal{O}\left(\left(\frac{\mu_{m+2}}{\mu_1}\right)^{k-1}\right) \\
&= \frac{\left(\sum_{j=2}^{m+1} f_j w_j\right)_i (\mu_2 - \mu_1)}{f_1 (w_1)_i} \left(\frac{\mu_2}{\mu_1} - 1\right) \left(\frac{\mu_2}{\mu_1}\right)^{k-2} + \mathcal{O}\left(\left(\frac{\mu_{m+2}}{\mu_1}\right)^{k-1}\right). \qquad (5.7)
\end{aligned}$$

Then the ratio of two consecutive differences is

$$\begin{aligned}
r^{(k)} := \frac{\widehat{\mu}_1^{(k)} - \widehat{\mu}_1^{(k-1)}}{\widehat{\mu}_1^{(k-1)} - \widehat{\mu}_1^{(k-2)}} &= \frac{\frac{\left(\sum_{j=2}^{m+1} f_j w_j\right)_i (\mu_2-\mu_1)}{f_1 (w_1)_i} \left(\frac{\mu_2}{\mu_1} - 1\right) \left(\frac{\mu_2}{\mu_1}\right)^{k-2} + \mathcal{O}\left(\left(\frac{\mu_{m+2}}{\mu_1}\right)^{k-1}\right)}{\frac{\left(\sum_{j=2}^{m+1} f_j w_j\right)_i (\mu_2-\mu_1)}{f_1 (w_1)_i} \left(\frac{\mu_2}{\mu_1} - 1\right) \left(\frac{\mu_2}{\mu_1}\right)^{k-3} + \mathcal{O}\left(\left(\frac{\mu_{m+2}}{\mu_1}\right)^{k-2}\right)} \\
&= \frac{\mu_2}{\mu_1} + \mathcal{O}\left(\left(\frac{\mu_{m+2}}{\mu_2}\right)^{k-3}\right). \qquad (5.8)
\end{aligned}$$

In the last equation, we have assumed that not all of $f_2, f_3, \ldots, f_{m+1}$ are zero. Otherwise, our methodology approximates $\frac{\mu_{m+2}}{\mu_1}$, which can serve the same purpose in deciding whether $x^{(k_0)}$ being positive is the result of round-off error or not.

We consider the ratios $r^{(k)}$ as estimates of $\frac{\mu_2}{\mu_1}$. In practice, we can perform a few more iterations beyond $k_0$ and compute consecutive values for $r^{(k)}$. When two consecutive terms differ in a prescribed small number of floating points (1 or 2), then we can use $r^{(k)}$ to decide if $x_0 \in X_A(\mathbb{R}^n_+)$ or not.

**Case** index$_{\mu_2}(B) = m > 1$**.**

Assume $\mu_2 \neq 0$, otherwise $x^{(k)}$ belongs to the Perron eigenspace for all $k \geq m$. Also, for simplicity of the presentation, we perform the analysis when $m = 2$ and the algebraic multiplicity of $\mu_2 \neq 0$ is also 2; the analysis for several Jordan blocks corresponding to $\mu_2$ and $m > 2$ is analogous. In this case, the vector $w_3$ is a generalized eigenvector of $\mu_2$. Then, relation (5.5) takes the form

$$B^k y_0 = f_1 \mu_1^k w_1 + \left[\left(f_2 + \frac{f_3}{\mu_2} k\right) w_2 + f_3 w_3\right] \mu_2^k + \mathcal{O}(\lambda_4^k) w, \tag{5.9}$$

where $w$ belongs to the generalized eigenspace of $\mu_4$. As before,

$$\begin{aligned} r^{(k)} &= \frac{\widehat{\mu}_1^{(k)} - \widehat{\mu}_1^{(k-1)}}{\widehat{\mu}_1^{(k-1)} - \widehat{\mu}_1^{(k-2)}} \\ &= \frac{\left[\left(f_2\mu_2 + f_3(k-1)\right)(w_2)_i + f_3\mu_2(w_3)_i\right]\left(\frac{\mu_2}{\mu_1} - 1\right) + f_3(w_2)_i\left(\frac{\mu_2}{\mu_1}\right)}{\left[\left(f_2\mu_2 + f_3(k-2)\right)(w_2)_i + f_3\mu_2(f_3)_i\right]\left(\frac{\mu_2}{\mu_1} - 1\right) + f_3(w_2)_i\left(\frac{\mu_2}{\mu_1}\right)} \frac{\mu_2}{\mu_1} \\ &\quad + \mathcal{O}\left(\left(\frac{\mu_3}{\lambda_2}\right)^{k-3}\right). \end{aligned} \tag{5.10}$$

We can rewrite (5.10) as

$$r^{(k)} = \frac{\widehat{\mu}_1^{(k)} - \widehat{\mu}_1^{(k-1)}}{\widehat{\mu}_1^{(k-1)} - \widehat{\mu}_1^{(k-2)}} = \frac{f_1' + f_2'(k-1)}{f_1' + f_2'(k-2)} \frac{\mu_2}{\mu_1} + \mathcal{O}\left(\left(\frac{\mu_3}{\mu_2}\right)^{k-3}\right). \tag{5.11}$$

The coefficient of $\frac{\mu_2}{\mu_1}$ is not 1 but tends to 1 as $k$ tends to infinity. This means that the convergence is slower in this case. However, $r^{(k)}$ can still play the same decisive role as in the previous case.

We now give an algorithm (in pseudocode) that implements our analysis above:

**Algorithm** (deciding membership in $X_A(\mathbb{R}^n_+)$)

```
Input:  A ∈ ℝ^{n×n} (irreducible essentially nonnegative), x^{(0)} ∈ ℝ^n, ε
% ε is the desired precision for the estimate μ_2/μ_1 (e.g., 10^{-2})
% tol below denotes the machine precision (e.g., eps in Matlab)

compute h(A) = sup{h | min_j(1 + a_jj) > 0}
determine h ∈ (0, h(A)) such that B = I + hA is invertible
compute x^{(k)} = B x^{(k-1)}  (k = 1, 2, ...)
until x^{(k)} is positive or negative at k = k_0.
if x^{(k_0)} < 0 then
        Output:  x^{(0)} ∉ X_A(ℝ^n_+); stop
else
        % Check if positivity is due to round-off error
        reset x^{(0)} = x^{(k_0)}; i = index of max_j ((x^{(0)})_j)
        compute x^{(k)} = B x^{(k-1)}  (k = 1, 2, ...)
```

$$\widehat{\mu}_1^{(k)} = \frac{(x^{(k)})_i}{(x^{(k-1)})_i}$$

$$r^{(k)} = \frac{\widehat{\mu}_1^{(k)} - \widehat{\mu}_1^{(k-1)}}{\widehat{\mu}_1^{(k-1)} - \widehat{\mu}_1^{(k-2)}}$$

```
        until (|r^{(k)} - r^{(k-1)}| ≤ ε |r^{(k)}|);
        if (r^{(k)})^{k_0} ~ tol then
                Output:  x^{(0)} ∉ X_A(ℝ^n_+); stop;
        else
                Output:  x^{(0)} ∈ X_A(ℝ^n_+); stop;
        end if
end if
end
```

# 6 Numerical illustration

In this section we apply the iterative test proved in Theorem 5.1 and implemented in the Algorithm of Section 5.

**Example 6.1** Consider the irreducible and essentially nonnegative matrix

$$A = \begin{pmatrix} 1/3 & 3 & 2 \\ 2 & -1/4 & 1 \\ 3 & 3 & 1/3 \end{pmatrix}.$$

Notice that $h(A) = 4$ so we must choose $h \in (0, 4)$ so that $B = I + hA$ is invertible. Since the eigenvalues of $A$ are $\lambda_1 = 4.6047, \lambda_2 = -2.4890$, and $\lambda_3 = -1.6990$, it follows that $B$ is invertible for all $h \in (0, h(A))$ except for $h = 1/2.4890$ and $h = 1/1.6990$. We set $h = 3$ and work with the nonnegative irreducible matrix

$$B = I + hA = \begin{pmatrix} 2 & 9 & 6 \\ 6 & 1/4 & 3 \\ 9 & 9 & 2 \end{pmatrix}.$$

Observe that the eigenvalues of $B$ are $\mu_1 = 14.8140 > |\mu_2| = |-6.4671| > |\mu_3| = |-4.0970|$ and so the matrix $B$ has, as predicted, a sole dominant eigenvalue.

We wish to consider all three cases distinguished in Theorem 5.1 and the Algorithm. To begin with, we select the initial point to be $x_0 = (0 \ \ -1 \ \ 10)^T$. Proceeding with our scheme, we have

$$x^{(1)} = Bx_0 = \begin{pmatrix} 51 \\ 29.75 \\ 11 \end{pmatrix} \in \mathbb{R}^3_+,$$

i.e., in just one iteration, the sequence $x^{(k)}$ becomes positive and thus we probably have $x_0 \in X_A(\mathbb{R}^3_+)$. Although this situation is clear, we follow the remaining steps of the Algorithm to illustrate how the rest of it works.

We apply the power method with initial vector $x^{(1)}$ normalized with respect to $\| \cdot \|_\infty$. We then compute $r^{(k)}$, $k = 1, 2, \ldots$, in order to estimate the ratio $\frac{\mu_2}{\mu_1}$. After 9 iterations we observe convergence of $r^{(k)}$ in two floating points; that is, $r^{(9)} \approx -0.44$ which is far from machine tolerance. Thus, it is confirmed that $x_0 \in X_A(\mathbb{R}^3_+)$. We notice that since our concern lies primarily with a qualitative description of entries of $x^{(k)}$ (and not convergence), we do not scale the iterate vectors.

Next, let us consider the initial vector $x_0 = (0 \ \ -80 \ \ 12)^T$. Applying $B$ to $x_0$ only twice, we obtain an entrywise negative vector:

$$x^{(1)} = Bx_0 = \begin{pmatrix} -648 \\ 16 \\ -696 \end{pmatrix}, \quad x^{(2)} = Bx^{(1)} = \begin{pmatrix} -5328 \\ -5972 \\ -7080 \end{pmatrix}.$$

It follows, as a consequence of Theorem 5.1 and the Algorithm, that $x \notin X_A(\mathbb{R}^n_+) \cup \mathcal{L}_B$.

Next, we select $x_0 = w_2 + w_3$, where $w_2$ and $w_3$ are eigenvectors corresponding to $\mu_2$ and $\mu_3$, respectively; that is, $x_0$ is an element of $\mathcal{L}_B$. Our scheme produces the expected result: For all iterates until the 44-th, vectors $x^{(k)}$ have positive and negative components that vary predictably based on whether the power of $B$ is odd or even. Specifically, $(x^{(k)})_1 > 0$ and $(x^{(k)})_{2,3} < 0$ if the power is odd; $(x^{(k)})_1 < 0$ and $(x^{(k)})_{2,3} > 0$ if the power is even:

$$x^{(1)} = \begin{pmatrix} 5.1022 \\ -1.5762 \\ -4.5204 \end{pmatrix}, \quad x^{(2)} = \begin{pmatrix} -32.7236 \\ 16.6129 \\ 21.0734 \end{pmatrix}, \text{ etc.}$$

However, because of the introduction of a direction along $w_1$ due to round-off error, in the 44-th step, the iterate becomes positive.

The Algorithm then applies the power method with initial vector the normalization of $x^{(44)}$ to estimate the ratio $\frac{\mu_2}{\mu_1}$. In the 7-th iteration of the power method, convergence is achieved in two floating points: $r^{(7)} \approx -0.44$. Then, $(-0.44)^{44} = 2.0508e - 016$ that is of the order of the machine tolerance (working with Matlab). Hence we have $x_0 \in \mathcal{L}_B$.

Lastly, we apply the same procedure to $x_0 = w_2 + w_3$ with $w_2$ and $w_3$ truncated to four decimal points. That is, $x_0 = (-0.7992 \;\; 0.0732 \;\; 1.0070)^T$. We find a positive vector for the first time at the 13-th iteration and estimate the ratio in the 8-th iteration of the power method in two floating points to be $r^{(8)} = -0.44$. As $(-0.44)^{13} = -2.3168e - 005$, which is far from machine tolerance, we conclude that $x_0 \in X_A(\mathbb{R}^3_+)$. This was indeed expected because the truncation of $w_2$ and $w_3$ we used introduced a direction along $w_1$.

Since our test typically provides an entrywise positive or negative vector in just a few iterations, given a vector that produces a predictable sign pattern variation for a large number of iterates, is a good warning sign that this vector is not an element of $X_A(\mathbb{R}^n_+)$.

**Example 6.2** We consider the irreducible and essentially nonnegative matrix

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 5/3 & -1/3 & 2/3 \\ 1/3 & 4/3 & 1/3 \end{pmatrix}.$$

Since $h(A) = 3$, we must choose $h \in (0, 3)$ so that $B = I + hA$ is invertible. As the eigenvalues of $A$ are $\lambda_1 = 2, \lambda_2 = \lambda_3 = -1$, it follows that $B$ is invertible for all $h \in (0, h(A))$

except $h = 1$. We set $h = 2$ and work with the nonnegative irreducible matrix

$$B = I + hA = \begin{pmatrix} 1 & 2 & 2 \\ 10/3 & 1/3 & 4/3 \\ 2/3 & 8/3 & 5/3 \end{pmatrix}.$$

Observe that the eigenvalues of $B$ are $\mu_1 = 5$ and $\mu_2 = \mu_3 = -1$ with $\text{index}_{\mu_2}(B) = 2$.

We first choose $x_0 = (-1 \;\; 2 \;\; 2)^T$ and find $x^{(1)} = Ax_0 > 0$. That is, in just one iteration we are able to determine that $x_0 \in X_A(\mathbb{R}^3_+)$. Our procedure for the determination of $\frac{\mu_2}{\mu_1}$ converges in the 9-th iteration to $-0.23$.

Next, we select $x_0 = w_2 + w_3 = (-1 \quad 0 \quad 1)^T$, where $w_2$ and $w_3$ are the eigenvector and the generalized eigenvector of $\mu_2$, respectively; that is, $x_0 \in \mathcal{L}_B$. We get that $x^{(k)}$ becomes positive for the first time at $k = 27$. The application of the second part of our procedure estimates the ratio $r^{(k)}$ in two floating points at $k = 6$ to be $r^{(6)} = -0.21$. Since $(-0.21)^{27} = -5.0110e - 019$, it is confirmed that $x_0 \in \mathcal{L}_B$. We remark that although $\text{index}_{\mu_2}(B) = 2$, the second part of our procedure converges fast (6 iterations) as compared to the slower behavior of the first part (27 iterations).

## 7 The general case

When $A \overset{\text{e}}{\geq} 0$ is possibly reducible, an algorithmic characterization of $X_A(\mathbb{R}^n_+)$ via $B = I + hA$ is still attainable, because it remains true that $X_A(\mathbb{R}^n_+) = X_{A,h}(\mathbb{R}^n_+)$. However, the development of an algorithm is complicated by several factors:

- The algebraic multiplicity, geometric multiplicity and index of $\rho(B)$ can all be greater than one.
- Eigenvectors corresponding to $\rho(B)$ can be taken to be nonnegative but not necessarily positive.
- There may be nonnegative eigenvectors corresponding to non-dominant eigenvalues of $B$.

As a consequence of the above complications, it is possible but more challenging to design a “black box” algorithm for membership in $X_A(\mathbb{R}^n_+)$ in the general case. Instead, we choose to describe the situation in all possible cases.

Before we do so, we need to recall some definitions. We consider $B$ in Frobenius normal form with the vertices in its directed graph partitioned in equivalence classes. A class is called *basic* if the corresponding block in the Frobenius normal form has spectral radius equal to $\rho(B)$. We call a class *final* if no other class has access to it in the reduced graph of $B$. For details, see [2, Chapter 2, Section 3].

By Lemma 4.3 we can assume that the matrix $B$ has $p$ distinct eigenvalues, ordered as

$$\rho(B) = \mu_1 > |\mu_2| \geq |\mu_3| \geq \cdots \geq |\mu_p|,$$

with multiplicities $m_1, m_2, \ldots, m_p$, respectively. Let also $w_1, w_2, \ldots, w_n$ be a Jordan basis for $\mathbb{C}^n$ consisting of (generalized) eigenvectors of $B$. We may take this basis so that all vectors $w_1, w_2, \ldots, w_{m_1}$ are nonnegative; see e.g., [2, Chapter 2, Theorem (3.20)]. It follows that

$$\mathrm{Span}\{w_{m_1+1}, w_{m_1+2} \ldots, w_n\} \;=\; \mathcal{L}_B = \bigoplus_{\mu \neq \rho(B)} \mathcal{N}_\mu(B).$$

Let us now consider an initial vector $x_0$ and iterates $x^{(k)}$ as in (5.1) and (5.2), respectively.

For convenience denote the coefficients in (5.1) associated with $\rho(B)$ by

$$\hat{\mathbf{c}} = (c_1 \; c_2 \; \cdots \; c_{m_1})^T.$$

<u>**Case 1:** $\mathrm{index}_{\rho(B)}(B) = 1$</u>

<u>If $\hat{\mathbf{c}} \neq 0$</u>, the sequence $\{x^{(k)}\}$ approaches the subspace spanned by $\{w_1, w_2, \ldots, w_{m_1}\}$ with convergence ratio $\left|\frac{\mu_2}{\mu_1}\right|$. In the reducible case, $x^{(k)}$ does not necessarily become positive or negative. In order to become positive, it must be that $\hat{\mathbf{c}} > 0$ and all the final classes in the Frobenius normal form of $B$ must be basic classes; see [2, Theorem (3.10)].

<u>If $\hat{\mathbf{c}} \geq 0$, $\hat{\mathbf{c}} \neq 0$</u>, there can be entries of $x^{(k)}$ that tend to zero. These entries correspond to the classes that are final and not basic. If for large enough $k$ all the entries that tend to zero are positive, then $x_0 \in X_A(\mathbb{R}^n_+)$. If there exists an entry which is negative or changes its sign periodically, then $x_0 \notin X_A(\mathbb{R}^n_+)$.

<u>If $\hat{\mathbf{c}} \ngeq 0$</u>, $x^{(k)}$ will not become a nonnegative vector in the limit. In this case, round-off errors do not affect the situation.

As a consequence of the above observations, a numerical test to decide whether $x_0 \in X_A(\mathbb{R}^n_+)$ or not would have to be along the following lines:

**1.** If $\{x^{(k)}\}$ becomes a positive vector, we conclude that $x_0 \in X_A(\mathbb{R}^n_+)$.
Otherwise,

**2.** Check if the entries tending to zero are positive or not. The ratio of convergence is at most $\left|\frac{\mu_2}{\mu_1}\right|$. Hence, for sufficient large $k$, any entry that tends to zero becomes smaller than *tol* in absolute value. If it is negative at this instant, we conclude that $x_0 \notin X_A(\mathbb{R}^n_+)$. If it is positive, we replace it by zero (which is its limit) and continue the iterations of the power method. In this way, at some iterate $k$, all the entries tending to zero will have been replaced by zero.

**3.** Consider the entries that correspond to the basic classes. We apply the procedure used in the irreducible case in order to estimate the ratio $\frac{\mu_2}{\mu_1}$. If $\{x^{(k)}\}$ becomes nonnegative, we check if this is due to round-off errors as in the irreducible case. If some negative entries persist until $\left|\frac{\mu_2}{\mu_1}\right|^k < tol$, then we conclude $x_0 \notin X_A(\mathbb{R}^n_+)$.

<u>If $\hat{\mathbf{c}} = 0$</u>, then, theoretically, $x^{(k)} \in \mathcal{L}_B$ $(k = 0, 1, \ldots)$. The power method will converge to the sum of the eigenspaces corresponding to the most dominant eigenvalue present in (5.1). Since, $B$ is reducible, an eigenvalue $\mu_j$, $(j \geq 2)$ may have a nonnegative eigenvector. Thus it is possible that $x_0 \in X_A(\mathbb{R}^n_+)$. To decide, we can apply the numerical test described above for when $\hat{\mathbf{c}} \neq 0$.

<u>**Case 2:** $\mathrm{index}_{\rho(B)}(B) > 1$</u>

In this case, the behavior of the power method is quite different. We suppose first that there exists only one block in the Jordan canonical form corresponding to $\mu_1$, so $s = m_1$. Then,

$$\begin{aligned}
x^{(k)} &= B^k x_0 = c_1 B^k w_1 + c_2 B^k w_2 + \cdots + c_s B^k w_s + \cdots + c_n B^k w_n \\
&= c_1 {\mu_1}^k w_1 + c_2 \left( {\mu_1}^k w_2 + \begin{pmatrix} k \\ 1 \end{pmatrix} \mu_1^{k-1} w_1 \right) + \cdots \\
&+ c_s \left( {\mu_1}^k w_s + \begin{pmatrix} k \\ 1 \end{pmatrix} \mu_1^{k-1} w_{s-1} + \cdots + \begin{pmatrix} k \\ s-1 \end{pmatrix} \mu_1^{k-s+1} w_1 \right) \\
&+ c_{m_1+1} {\mu_2}^k w_{m_1+1} + \cdots + c_n {\mu_p}^k w_n \quad (k = 0, 1, \ldots).
\end{aligned} \tag{7.1}$$

Clearly, the power method converges to the dominant eigenvector $w_1$, but the convergence is very slow. If we normalize $x^{(k)}$ by dividing by $\begin{pmatrix} k \\ s-1 \end{pmatrix} \mu_1^{k-s+1}$, it converges like $\frac{1}{k}$. This converges occurs for entries of $x^{(k)}$ for which some generalized eigenvectors have corresponding positive values. The ratio of convergence of the entries of $x^{(k)}$ corresponding to non-basic classes is also at most $\left|\frac{\mu_2}{\mu_1}\right|$. Thus, they become smaller than *tol* quickly and will be replaced by zeros. The remaining nonzero entries correspond to the basic classes in the Jordan canonical form. It is easily seen that the dominant term in (7.1) is

$$c_s \left( {\mu_1}^k w_s + \begin{pmatrix} k \\ 1 \end{pmatrix} \mu_1^{k-1} w_{s-1} + \cdots + \begin{pmatrix} k \\ s-1 \end{pmatrix} \mu_1^{k-s+1} w_1 \right) \tag{7.2}$$

and thus the normalized limit of $x^{(k)}$ is $c_s w_1$. This means that the common sign of all the entries of $x^{(k)}$ corresponding to the nonzero entries of the vector in (7.2) is determined by the sign of $c_s$. If $c_s = 0$, then the sign is determined by $c_{s-1}$ and so on. In the case of many Jordan blocks, for each block there corresponds a set of such entries, whose sign is determined by the coefficient of the highest generalized eigenvector in the Jordan chain. To decide if $x_0 \in X_A(\mathbb{R}^n_+)$ or not, we can adapt and apply the numerical test found in Case 1. The step of determining the sign of the entries tending to zero works in the same way. In the step of estimating the ratio $\frac{\mu_2}{\mu_1}$, since the convergence is like $\frac{1}{k}$, it can be shown that the quotient $r(k)$ tends to 1 like the $\frac{k-s-2}{k-s}$. This informs us that we are in the reducible case with $\mathrm{index}_{\mu_1}(B) > 1$. The consequence is that a determination of whether $x_0 \in X_A(\mathbb{R}^n_+)$ can indeed be made, however, it may require a large number of iterations.

## References


[1] A. Berman, M. Neumann, and R.J. Stern. *Nonnegative Matrices in Dynamic Systems*. Wiley-Interscience, 1989.

[2] A. Berman and B. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. SIAM, Philadelphia, 1994.

[3] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

[4] M. Neumann and R.J. Stern. Boundary results for positively invariant cones and their reachability cones. *Linear and Multilinear Algebra*, 17:143–154, 1985.

[5] M. Neumann and R.J. Stern. Cone reachability for linear differential systems. *Applicable Analysis*, 20:57–71, 1986.

[6] M. Neumann, R.J. Stern, and M. Tsatsomeros. The reachability cones of essentially nonnegative matrices. *Linear and Multilinear Algebra*, 28:213–224, 1991.

[7] D. Noutsos and M. Tsatsomeros. Reachability and holdability of nonnegative states. *SIAM Journal on Matrix Analysis and Appplications*, 30(2):700–712, 2008.

[8] R.T. Rockafellar. *Convex Analysis*. Princeton University Press, New Jersey, 1997.

[9] R.J. Stern. A note on positively invariant cones. *Applied Mathematics and Optimization*, 9:67–72, 1982.

[10] D. Watkins. *Fundamentals of Matrix Computations*, Second ed. Wiley-Interscience, New York, 2002.