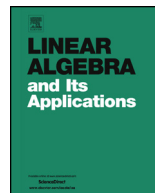




Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa



On the joint spectral radius of nonnegative matrices

Vuong Bui¹

Institut für Informatik, Freie Universität Berlin, Takustraße 9, 14195 Berlin, Germany



ARTICLE INFO

Article history:

Received 20 December 2021

Accepted 27 August 2022

Available online 1 September 2022

Submitted by M. Tsatsomeros

MSC:

15A18

15A60

65F15

Keywords:

Joint spectral radius

Nonnegative matrices

Fekete's lemma

Joint spectral radius theorem

ABSTRACT

We give an effective bound of the joint spectral radius $\rho(\Sigma)$ for a finite set Σ of nonnegative matrices: For every n ,

$$\sqrt[n]{\left(\frac{V}{UD}\right)^D \max_C \max_{i,j \in C} \max_{A_1, \dots, A_n \in \Sigma} (A_1 \dots A_n)_{i,j}} \leq \sqrt[n]{D \max_C \max_{i,j \in C} \max_{A_1, \dots, A_n \in \Sigma} (A_1 \dots A_n)_{i,j}},$$

where $D \times D$ is the dimension of the matrices, U, V are respectively the largest entry and the smallest entry over all the positive entries of the matrices in Σ , and C is taken over all strongly connected components in the dependency graph. The dependency graph is a directed graph where the vertices are the dimensions and there is an edge from i to j if and only if $A_{i,j} \neq 0$ for some matrix $A \in \Sigma$.

Furthermore, a bound on the norm is also given: If $\rho(\Sigma) > 0$ then there exist a nonnegative integer r and two positive numbers α, β so that for every n ,

$$\alpha n^r \rho(\Sigma)^n \leq \max_{A_1, \dots, A_n \in \Sigma} \|A_1 \dots A_n\| \leq \beta n^r \rho(\Sigma)^n.$$

Corollaries of the approach include a simple proof for the joint spectral theorem for finite sets of nonnegative matrices and the convergence rate of some sequences. The method in use is

E-mail address: bui.vuong@yandex.ru.

¹ The author is supported by the Deutsche Forschungsgemeinschaft (DFG) Graduiertenkolleg “Facets of Complexity” (GRK 2434).

mostly based on Fekete's lemma, for both submultiplicative and supermultiplicative sequences.

© 2022 Elsevier Inc. All rights reserved.

1. Introduction

Joint spectral radius is a generalization of spectral radius to a set of matrices, which was firstly introduced in [1] by Rota and Strang. This has caught a lot of attention with its theoretical interest as well as its applications in engineering fields. We advise the readers to check [2] for a book with a comprehensive treatment of the subject. As for application purposes, we naturally need a method to estimate the joint spectral radius. In this section, we begin with the definition of the radius, followed by some known methods for estimation, and conclude with our proposed bounds for *finite* sets of *nonnegative matrices*. Our method mainly uses Fekete's lemma [3] for both submultiplicative and supermultiplicative sequences.

Given a finite set Σ of square matrices in $\mathbb{C}^{d \times d}$, we denote

$$\|\Sigma^n\| = \max_{A_1, \dots, A_n \in \Sigma} \|A_1 \dots A_n\|.$$

In [1], the joint spectral radius $\rho(\Sigma)$ of the set Σ is defined to be the limit

$$\rho(\Sigma) = \lim_{n \rightarrow \infty} \sqrt[n]{\|\Sigma^n\|}.$$

In fact, $\rho(\Sigma)$ is also defined for infinite bounded sets Σ . However, we only consider *finite* sets of matrices in this text. Whether the results hold for infinite sets or how they can be extended is left open.

The following result, which appears in most of the sources, is used to prove that the limit exists. We provide it again here as it also gives a bound on the radius. It is often expressed in submultiplicative norms and in a slightly different form. However, for convenience, the *maximum norm* will be used for a matrix throughout the work unless stated otherwise.² Also, to emphasize that the dimension is a constant, we denote $D = d$ hereafter.

Proposition 1 (*A slight modification of the popular proof for the maximum norm*³). *The following limit exists and can be expressed as:*

² The maximum norm is not a submultiplicative norm, as required in the definition of a matrix norm in some texts. However, it does not affect the asymptotic behavior. To be more precise, the maximum norm here is the largest absolute value of any entry in the matrix. It becomes the largest entry when only nonnegative matrices are considered.

³ In the form for a submultiplicative norm, the expression on the right would be $\inf_n \sqrt[n]{\|\Sigma^n\|}$.

$$\lim_{n \rightarrow \infty} \sqrt[n]{\|\Sigma^n\|} = \inf_n \sqrt[n]{D\|\Sigma^n\|}.$$

Proof. For any two matrices A, B , we have

$$\|AB\| = \max_{i,j} |(AB)_{i,j}| = \max_{i,j} \left| \sum_k A_{i,k} B_{k,j} \right| \leq D\|A\|\|B\|.$$

For any two positive integers m, n , we have

$$\|\Sigma^{m+n}\| = \|A_1 \dots A_{m+n}\| \leq D\|A_1 \dots A_m\|\|A_{m+1} \dots A_{m+n}\| \leq D\|\Sigma^m\|\|\Sigma^n\|,$$

where A_1, \dots, A_{m+n} are some matrices from Σ .

Writing differently, $D\|\Sigma^{m+n}\| \leq (D\|\Sigma^m\|)(D\|\Sigma^n\|)$ means the sequence $\{D\|\Sigma^n\|\}_n$ is submultiplicative. By Fekete's lemma, $\sqrt[n]{D\|\Sigma^n\|}$ converges to $\inf_n \sqrt[n]{D\|\Sigma^n\|}$, which is also the limit of $\sqrt[n]{\|\Sigma^n\|}$. \square

One of the popular ways to estimate the joint spectral radius is as follows (see Proposition 1.6 and Section 2.3.3 on “Branch and Bounds Methods” in the book [2]). For any sequence of matrices $A_1, \dots, A_m \in \Sigma$, the m -th root of the (ordinary) spectral radius of $A_1 \dots A_m$ is a lower bound for $\rho(\Sigma)$. Denote⁴

$$P_m(\Sigma) = \max_{A_1, \dots, A_m \in \Sigma} \rho(A_1 \dots A_m).$$

Together with the bound from Proposition 1, we can bound $\rho(\Sigma)$ from both sides: For any m ,

$$\sqrt[m]{P_m(\Sigma)} \leq \rho(\Sigma) \leq \sqrt[m]{D\|\Sigma^m\|}. \quad (1)$$

One of the points supporting this method of bounding is that the limit superior of the sequence for the left side and the limit of the sequence for the right side (with respect to m) are equal to $\rho(\Sigma)$. In fact, $\limsup_{n \rightarrow \infty} \sqrt[n]{P_n(\Sigma)}$ is called the *generalized spectral radius* of Σ . That the two radii are equal for the case of *finite* sets Σ is the content of the joint spectral radius theorem. It was conjectured by Daubechies and Lagarias in [4] and gets proved in [5] for the first time by Berger and Wang. The readers can also see Theorem 2.3 in the book [2] for a reference.

Theorem 1 (The joint spectral radius theorem [4] [5]). *For every finite set Σ of matrices,*

$$\limsup_{n \rightarrow \infty} \sqrt[n]{P_n(\Sigma)} = \lim_{n \rightarrow \infty} \sqrt[n]{\|\Sigma^n\|}.$$

⁴ Note that when we write $\rho(A)$ for a matrix A , we mean the classic spectral radius for A . There should be no confusion as the type of spectral radius is decided by the argument for ρ .

As discussed in [6, Section 2], the limit superior is shown to be not replaceable by a limit in general with a counterexample. Also, it is asked there that: What is the convergence rate of $\max_{1 \leq m \leq n} \sqrt[m]{P_m(\Sigma)}$ and $\min_{1 \leq m \leq n} \sqrt[m]{D\|\Sigma^m\|}$ (with respect to n)⁵ to $\rho(\Sigma)$? This question is critical to the efficiency of the bound in (1). Section 3 will show that both sequences converge at the rate $O(1/n)$ and give a simple proof of the joint spectral radius theorem, both however only for finite sets of nonnegative matrices.

The following bound in [7] has a clearer convergence rate: For every n ,

$$\sqrt[n]{f(n)\|\Sigma^n\|} \leq \rho(\Sigma) \leq \sqrt[n]{\|\Sigma^n\|}, \quad (2)$$

where $f(n)$ is rather complicated and it may grow very low (note that the norm in (2) is a submultiplicative norm). The work [7] describes $f(n)$ explicitly, but loosely speaking, $f(n)$ is in general roughly about C^{-n^α} for a constant C and $\alpha = \ln(D-1)/\ln D$. Although the bound in (2) is very interesting, it is hard to estimate $\rho(\Sigma)$ effectively as the ratio of the two bounds is large.

If one restricts the scope to *irreducible* sets of matrices, we have the following bound in [8]: There is constant γ so that for every n ,

$$\sqrt[n]{\gamma^{1+\ln n}\|\Sigma^n\|} \leq \rho(\Sigma) \leq \sqrt[n]{\|\Sigma^n\|}.$$

Note that the norm here is also submultiplicative. Also note that the quantity $\gamma^{1+\ln n}$ under the n -th root is actually of order n^{-t} for some t . While the condition of irreducible sets is algebraic, our condition given in Theorem 2 below is on the signs of the entries of matrices.

Before introducing our bounds, we need to define the dependency graph.

Definition 1. The *dependency graph* of a set of matrices Σ is a directed graph where the vertices are the dimensions and there is an edge from i to j if and only if $A_{i,j} \neq 0$ for some matrix $A \in \Sigma$ (loops are allowed). This graph can be partitioned into strongly connected components, for which we will call *components* for short. For a component C , we denote

$$\|\Sigma^n\|_C = \max_{A_1, \dots, A_n \in \Sigma} \max_{i,j \in C} |(A_1 \dots A_n)_{i,j}|.$$

Also, we denote by $\delta(i, j)$ the distance from i to j in the dependency graph.

Our proposed bound is given in the following theorem, which works only for *finite* sets of *nonnegative matrices* and will be proved in Section 2. From now on, we assume that at least one entry of some matrix of Σ is nonzero, as otherwise all matrices are

⁵ It was actually asked in [6] for $\sqrt[n]{\|\Sigma^n\|}$ for submultiplicative norms. We adapt it for the maximum norm.

zero matrices, which is trivial. The assumption is to let U, V in the theorem below well defined.

Theorem 2. *Given a finite set Σ of nonnegative matrices. For every n ,*

$$\sqrt[n]{\left(\frac{V}{UD}\right)^D \max_C \|\Sigma^n\|_C} \leq \rho(\Sigma) \leq \sqrt[n]{D \max_C \|\Sigma^n\|_C},$$

where $D \times D$ is the dimension of the matrices, U, V are respectively the largest entry and the smallest entry over all the positive entries of the matrices in Σ , and C is taken over all components in the dependency graph.

The merit of Theorem 2 is that the ratio between the upper bound and the lower bound is the n -th root of a constant, which means the interval length is $O(1/n)$. Therefore, when D and U/V are not too large, the gap between them can be reasonably small even with a not so large n . Although U/V can be arbitrarily large, the appearance of U and V is essential to the formula. For example, let Σ contain only one matrix A with its powers:

$$A = \begin{bmatrix} 1 & \frac{1}{N} \\ N & 1 \end{bmatrix}, \quad A^n = 2^{n-1} \begin{bmatrix} 1 & \frac{1}{N} \\ N & 1 \end{bmatrix} = 2^{n-1} A,$$

where N is a large number. The joint spectral radius is obviously 2 while $\max_C \|\Sigma^1\|_C = N$, therefore, the relation between U and V must present in the formula in some form.

Note that $\|\Sigma^n\|_C$ is still computed based on the computation of all $|\Sigma|^n$ combinations. It is not a bad approach as $\rho(\Sigma)$ is not approximable in polynomial time unless $P = NP$ (see [9]). Furthermore, the problem of checking $\rho(\Sigma) \leq 1$ is even undecidable (see [10]). On the other hand, the theorem applies very well if the set contains only one matrix, i.e. the case of the ordinary spectral radius. In fact, the bound (2) is perhaps the most effective estimation of that type for the spectral radius of a matrix so far in literature, while our result improves upon it for nonnegative matrices.

During the proof of Theorem 2, we also give the following bound on $\|\Sigma^n\|$ as in Theorem 3: If $\rho(\Sigma) > 0$ then there exists a nonnegative integer r so that for every n ,

$$\text{const } n^r \rho(\Sigma)^n \leq \|\Sigma^n\| \leq \text{const } n^r \rho(\Sigma)^n.$$

The inequalities show that $\sqrt[n]{\|\Sigma^n\|}$ converges to $\rho(\Sigma)$ at the rate $O(1/n)$.

Note that throughout the text, the notation const stands for some positive constant, and consecutive instances of const may present different constants.

In the theme of nonnegative matrices, we mention [11, Theorem 16], which works for nonnegative matrices only: Define a matrix S so that $S_{i,j} = \max_{A \in \Sigma} A_{i,j}$, we have

$$\frac{\rho(S)}{m} \leq \rho(\Sigma) \leq \rho(S). \quad (3)$$

The gap between the two bounds is obviously larger than that of Theorem 2. However, an advantage of this larger gap is that it asks for lower computational complexity when having to calculate the spectral radius of a single matrix. In fact, the authors of [11] asked whether the bound in (3) can be improved at a reasonable computational cost.

It is rather quite obvious that the method of bounding in Theorem 2 is asymptotically better than the other methods except that it is not clear for the one in (1). The comparison to the latter one can be done only in Section 3 when we have all the necessary results. In short, the method in Theorem 2 is better than the one in (1) by a root of a polynomial of degree r . However, there is a modification to make the two methods asymptotically equivalent.

2. Proof of Theorem 2

The central idea of the approach in Proposition 1 is $\|\Sigma^{m+n}\| \leq \text{const} \|\Sigma^m\| \|\Sigma^n\|$. The other direction $\|\Sigma^{m+n}\| \geq \text{const} \|\Sigma^m\| \|\Sigma^n\|$ suggests an alternative approach that works in the case of nonnegative matrices. While Proposition 1 gives an upper bound for $\rho(\Sigma)$, Proposition 2 below gives a lower bound.

Proposition 2. *Given a finite set Σ nonnegative matrices with a connected dependency graph, we have the following weak form of supermultiplicativity: For every m, n ,*

$$\|\Sigma^m\| \|\Sigma^n\| \leq \left(\frac{UD}{V} \right)^D \|\Sigma^{m+n}\|,$$

where U, V, D are defined as in Theorem 2.

Before presenting the proof, we first give a lemma that will be used in the manipulations of the proof.

Lemma 1. *Given some n matrices A_1, \dots, A_n from Σ , for each positive entry $(A_1 \dots A_n)_{i,j}$, we have*

$$V^n \leq (A_1 \dots A_n)_{i,j} \leq D^{n-1} U^n.$$

Proof. The entry of the product can be written as

$$(A_1 \dots A_n)_{i,j} = \sum_{\substack{k_0, k_1, \dots, k_n \\ k_0=i, k_n=j}} (A_1)_{k_0, k_1} (A_2)_{k_1, k_2} \dots (A_n)_{k_{n-1}, k_n}.$$

Each positive summand is obviously at least V^n and at most U^n . There are at most D^{n-1} summands. The conclusion follows. \square

Proof of Proposition 2. Let $\delta(i, j)$ be the distance from i to j in the dependency graph, we have

$$\begin{aligned}
 \|\Sigma^m\| \|\Sigma^n\| &= (A_1 \dots A_m)_{i,j} (A'_1 \dots A'_n)_{i',j'} \\
 &\leq V^{-\delta(j,i')} (A_1 \dots A_m)_{i,j} (B_1 \dots B_{\delta(j,i')})_{j,i'} (A'_1 \dots A'_n)_{i',j'} \\
 &\leq V^{-\delta(j,i')} (A_1 \dots A_m B_1 \dots B_{\delta(j,i')})_{i,j} (A'_1 \dots A'_n)_{i',j'} \\
 &\leq V^{-\delta(j,i')} \|\Sigma^{m+n+\delta(j,i')}\| \\
 &\leq V^{-\delta(j,i')} D \|\Sigma^{m+n}\| \|\Sigma^{\delta(j,i')}\| \\
 &\leq V^{-\delta(j,i')} D \|\Sigma^{m+n}\| D^{\delta(j,i')-1} U^{\delta(j,i')} \\
 &\leq \left(\frac{UD}{V}\right)^D \|\Sigma^{m+n}\|,
 \end{aligned}$$

where the matrices $A_1, \dots, A_m, A'_1, \dots, A'_n$ and the indices i, j, i', j' can be chosen from context. The matrices $B_1, \dots, B_{\delta(j,i')}$ are chosen from Σ so that $(B_1 \dots B_{\delta(j,i')})_{j,i'} > 0$. \square

Let $K = (\frac{V}{UD})^D$, we rewrite the inequality as

$$\|\Sigma^{m+n}\| \geq K \|\Sigma^m\| \|\Sigma^n\|.$$

It means $K \|\Sigma^{m+n}\| \geq (K \|\Sigma^m\|)(K \|\Sigma^n\|)$, i.e. the sequence $K \|\Sigma^n\|$ is supermultiplicative. As this sequence is also positive, $\sqrt[n]{K \|\Sigma^n\|}$ converges to $\sup_n \sqrt[n]{K \|\Sigma^n\|}$ by Fekete's lemma. This is also the limit of $\sqrt[n]{\|\Sigma^n\|}$.

We have now an effective bound of $\rho(\Sigma)$ for any set Σ of nonnegative matrices with a connected dependency graph.

Corollary 1. *If the dependency graph is connected, then for every n ,*

$$\sqrt[n]{\left(\frac{V}{UD}\right)^D \|\Sigma^n\|} \leq \rho(\Sigma) \leq \sqrt[n]{D \|\Sigma^n\|},$$

where U, V, D are defined as in Theorem 2.

When the concern is the value of $\|\Sigma^n\|$, we have

$$\text{const } \rho(\Sigma)^n \leq \|\Sigma^n\| \leq \text{const } \rho(\Sigma)^n. \quad (4)$$

We can extend the treatment to any set of nonnegative matrices, which do not necessarily have a connected dependency graph.

Consider the case when the dependency graph is not connected. For each component C , by restricting Σ to the indices in C , we have actually verified the limit

$$\lambda_C = \lim_{n \rightarrow \infty} \sqrt[n]{\|\Sigma^n\|_C},$$

where $\|\Sigma^n\|_C$ is defined as in Definition 1. Note that although a component of only one vertex without any loop is not considered as a connected graph, $\lambda_C = 0$ is still defined for such a component as $\|\Sigma^n\|_C = 0$ for any n . The formula in Corollary 1 still trivially applies to such a component.

The formula corresponding to (4) is

$$\text{const}(\lambda_C)^n \leq \|\Sigma^n\|_C \leq \text{const}(\lambda_C)^n.$$

Let $\lambda = \max_C \lambda_C$ where the maximum is taken over all the components C , we have

$$\text{const} \lambda^n \leq \max_C \|\Sigma^n\|_C \leq \text{const} \lambda^n. \quad (5)$$

When the concern is the maximum over all entries, we have the following theorem, which demonstrates the rate of convergence for $\sqrt[n]{\|\Sigma^n\|}$ and describes the value of $\|\Sigma^n\|$ better than the mere limit does when the limit is 1.

Theorem 3. *If $\lambda > 0$ then there exists a nonnegative integer r so that for every n ,*

$$\text{const} n^r \lambda^n \leq \|\Sigma^n\| \leq \text{const} n^r \lambda^n.$$

Note that when $\lambda = 0$, the bound still works for n large enough since there are only finitely many n so that $\|\Sigma^n\| > 0$.

Proof. We prove with the assumption that $\lambda > 0$.

For any i, j , the computation of $(A_1 \dots A_n)_{i,j}$ is usually written as

$$(A_1 \dots A_n)_{i,j} = \sum_{\substack{k_0, k_1, \dots, k_n \\ k_0 = i, k_n = j}} (A_1)_{k_0, k_1} (A_2)_{k_1, k_2} \dots (A_n)_{k_{n-1}, k_n}.$$

Partitioning a sequence k_0, k_1, \dots, k_n into vertices of the same components, we obtain

$$\begin{aligned} (A_1 \dots A_n)_{i,j} = & \sum_{\ell} \sum_{C_1, \dots, C_{\ell}} \sum_{m_1, \dots, m_{\ell}} \sum_{i_1, j_1, \dots, i_{\ell}, j_{\ell}} (A_1 \dots A_{m_1})_{i_1, j_1} (A_{m_1+1})_{j_1, i_2} (A_{m_1+2} \dots A_{m_1+m_2+1})_{i_2, j_2} \dots \\ & \dots (A_{n-m_{\ell}-m_{\ell-1}} \dots A_{n-m_{\ell}-1})_{i_{\ell-1}, j_{\ell-1}} (A_{n-m_{\ell}})_{j_{\ell-1}, i_{\ell}} (A_{n-m_{\ell}+1} \dots A_n)_{i_{\ell}, j_{\ell}}, \end{aligned} \quad (6)$$

where the sum is taken over all possible choices of: some number ℓ of components to consider, some different components C_1, \dots, C_ℓ , some partition of nonnegative parts $m_1 + \dots + m_\ell = n - \ell + 1$ and some indices $i_1, j_1 \in C_1, \dots, i_\ell, j_\ell \in C_\ell$ with $i_1 = i, j_\ell = j$ (note that if $m_t = 0$ then $i_t = j_t$).

To let the summand of (6) be positive, the sequence C_1, \dots, C_ℓ should form a chain in the sense that there is an edge uv with $u \in C_i, v \in C_{i+1}$ for any two adjacent C_i, C_{i+1} . Let r be one less than the maximal number of components C in a chain with $\lambda_C = \lambda$ over all the chains.

The right hand side of (6) is actually at most

$$\begin{aligned} & \sum_{\ell} \sum_{C_1, \dots, C_\ell} \sum_{m'} \sum_{\substack{\{m_t: 0 < \lambda_{C_t} < \lambda\} \\ \sum m_t = m'}} \sum_{\substack{\{m_t: \lambda_{C_t} = \lambda\} \\ \sum m_t = n - m' - \ell + 1}} \text{const} \left(\prod_{t: 0 < \lambda_{C_t} < \lambda} \|\Sigma^{m_t}\|_{C_t} \right) \\ & \quad \times \left(\prod_{t: \lambda_{C_t} = \lambda} \|\Sigma^{m_t}\|_{C_t} \right) \\ & \leq \sum_{\ell} \sum_{C_1, \dots, C_\ell} \sum_{m'} \text{const} (m')^{\ell - (r' + 1) - 1} (n - m' - \ell + 1)^{r'} (\lambda')^{m'} \lambda^{n - m' - \ell + 1} \\ & \leq \text{const } n^r \lambda^n, \end{aligned}$$

where r' is one less than the number of components C in the chain C_1, \dots, C_ℓ so that $\lambda_C = \lambda$, and λ' is some positive number less than λ so that if $\lambda_{C_t} < \lambda$ then $\lambda_{C_t} < \lambda'$. (The assumption $\lambda > 0$ is to make the products involving t not all empty.)

In total, we have $\|\Sigma^n\| \leq \text{const } n^r \lambda^n$. It remains to show the other direction that $\|\Sigma^n\| \geq \text{const } n^r \lambda^n$. Consider a chain C_1, \dots, C_ℓ with $r + 1$ components C attaining $\lambda_C = \lambda$.

Since $\|\Sigma^{n+\delta}\| \leq \text{const } \|\Sigma^n\| \|\Sigma^\delta\| = \text{const } \|\Sigma^n\|$ for any fixed δ , the following inequality holds for any fixed Δ :

$$\|\Sigma^n\| \geq \text{const} \sum_{\delta=0}^{\Delta} \|\Sigma^{n+\delta}\|.$$

If there is a path from i to j , we denote $B(i, j) = (A_1 \dots A_{\delta(i, j)})_{i, j}$ for some $A_1, \dots, A_{\delta(i, j)}$ from Σ so that $(A_1 \dots A_{\delta(i, j)})_{i, j} > 0$.

Let $\Delta = D^2$, we have

$$\begin{aligned} \sum_{\delta=0}^{\Delta} \|\Sigma^{n+\delta}\| & \geq \sum_{\substack{m_1 + \dots + m_\ell = n - \ell + 1 \\ m_t = 0 \text{ for } \lambda_{C_t} < \lambda}} \|\Sigma^{m_1}\|_{C_1} B(j_1, i_2) \|\Sigma^{m_2}\|_{C_2} \\ & \quad \dots \|\Sigma^{m_{\ell-1}}\|_{C_{\ell-1}} B(j_{\ell-1}, i_\ell) \|\Sigma^{m_\ell}\|_{C_\ell}, \end{aligned}$$

where two different values of each m_t for $\lambda_{C_t} = \lambda$ must be at least D apart, each pair i_t, j_t are so that $\|\Sigma^{m_t}\|_{C_t} = (A_1 \dots A_{m_t})_{i_t, j_t}$ for some $A_1, \dots, A_{m_t} \in \Sigma$ (for $m_t = 0$, we set $i_t = j_t = k$ for any element $k \in C_t$ and let $\|\Sigma^{m_t}\|_{C_t} = 1$).

Let S denote the right hand side. The reason for requiring any two values of each m_t to be at least D apart is to ensure no summand of S is counted twice as there are less than D matrices presenting in any $B(i, j)$. The reason for setting $\Delta = D^2$ is that the number of matrices presenting in each summand of S is larger than n by at most the number of components times the largest distance, which is always bounded by D^2 .

As we set $m_t = 0$ for $\lambda_{C_t} < \lambda$, each summand of S is at least $\text{const } \lambda^n$. Note that n^r is the order of the number of partitions $n_0 + \dots + n_r = n - \ell + 1$ even with the condition that for every i , the values of n_i in two different partitions are at least δ apart, for any fixed δ . Therefore, $S \geq \text{const } n^r \lambda^n$. The conclusion follows as $\|\Sigma^n\| \geq \text{const } S$. \square

Now we can see that Theorem 2 is a corollary of Theorem 3.

Proof of Theorem 2. If $\lambda = 0$, then $\rho(\Sigma) = 0$ and the bound in Theorem 2 trivially holds by (5). Suppose $\lambda > 0$. It follows from Theorem 3 that $\lambda = \rho(\Sigma)$. Since $\lambda = \max_C \lambda_C$, the bound in Corollary 1 should now become

$$\sqrt[n]{\left(\frac{V}{UD}\right)^D} \max_C \|\Sigma^n\|_C \leq \rho(\Sigma) \leq \sqrt[n]{D \max_C \|\Sigma^n\|_C}. \quad \square$$

3. Applications to the joint spectral radius theorem

In this section, we prove the joint spectral radius theorem for finite sets Σ of nonnegative matrices and deduce the convergence rates of some sequences. The readers should go through Section 2 first as the arguments below use some results from Section 2.

At first, for any connected component C and any m , we have

$$\begin{aligned} \|\Sigma^m\|_C &= (A_1 \dots A_m)_{i,j} \leq \text{const}(A_1 \dots A_m)_{i,j} (B_1 \dots B_{\delta(j,i)})_{j,i} \\ &\leq \text{const}(A_1 \dots A_m B_1 \dots B_{\delta(j,i)})_{i,i} \leq \text{const } \rho(A_1 \dots A_m B_1 \dots B_{\delta(j,i)}) \\ &\leq \text{const } P_{m+\delta(j,i)}(\Sigma), \end{aligned}$$

where $i, j \in C$ and A_1, \dots, A_m are given from context, and $B_1, \dots, B_{\delta(j,i)}$ are so that $(B_1 \dots B_{\delta(j,i)})_{j,i} > 0$. Note that if C is an unconnected component (containing only one vertex without loop), the inequality still holds trivially.

Taking over all components, we have

$$\max_C \|\Sigma^m\|_C \leq \max_{0 \leq \delta \leq D} \text{const } P_{m+\delta}(\Sigma).$$

Applying Theorem 2 to a set of only one matrix with $n = 1$, we have⁶

$$\begin{aligned} P_{m+\delta}(\Sigma) &= \max_{A_1, \dots, A_{m+\delta} \in \Sigma} \rho(A_1 \dots A_{m+\delta}) \leq \max_{A_1, \dots, A_{m+\delta} \in \Sigma} \text{const} \max_C \|\{A_1 \dots A_{m+\delta}\}\|_C \\ &= \text{const} \max_C \|\Sigma^{m+\delta}\|_C \leq \text{const} \max_C \|\Sigma^m\|, \end{aligned}$$

where the latest inequality is due to (5).

In total,

$$\text{const} \max_C \|\Sigma^m\|_C \leq \max_{0 \leq \delta \leq D} P_{m+\delta}(\Sigma) \leq \text{const} \max_C \|\Sigma^m\|_C.$$

It follows from (5) that

$$\text{const} \lambda^m \leq \max_{0 \leq \delta \leq D} P_{m+\delta}(\Sigma) \leq \text{const} \lambda^m. \quad (7)$$

Let $\tilde{P}_m(\Sigma) = \max_{0 \leq \delta \leq D} P_{m+\delta}(\Sigma)$. Although $\sqrt[n]{P_n(\Sigma)}$ does not necessarily converge, the sequence of $\sqrt[n]{\tilde{P}_n(\Sigma)}$ converges to λ . Together with Theorem 3, we have

$$\limsup_{n \rightarrow \infty} \sqrt[n]{P_n(\Sigma)} = \lim_{n \rightarrow \infty} \sqrt[n]{\|\Sigma^n\|},$$

which is the conclusion of the joint spectral radius theorem. (When Theorem 3 does not apply, i.e. $\lambda = 0$, the equality becomes trivial.)

Now we can see the convergence rates of some sequences that clarify the efficiency of the bound in (1).

At first, Theorem 3 gives

$$\text{const} m^r \lambda^m \leq \|\Sigma^m\| \leq \text{const} m^r \lambda^m,$$

which shows that the sequence $\min_{1 \leq m \leq n} \sqrt[n]{D \|\Sigma^m\|}$ converges at the rate $O(1/n)$ to λ .

The same convergence rate also applies to $\max_{1 \leq m \leq n} \sqrt[n]{P_m(\Sigma)}$. Indeed, it follows from (7) that

$$\lambda \min_{0 \leq \delta \leq D} \sqrt[m+\delta]{\text{const}} \leq \max_{0 \leq \delta \leq D} \sqrt[m+\delta]{P_{m+\delta}(\Sigma)} \leq \lambda \max_{0 \leq \delta \leq D} \sqrt[m+\delta]{\text{const}}.$$

It remains to compare the bound in (1) and the bound in Theorem 2. When $r = 0$, the two bounds are asymptotically equivalent. When $r > 0$, the ratio between the upper bound and the lower bound in (1) is at least the n -th root of a polynomial, which may reduce the efficiency when r is large. One may improve the traditional method in (1) by

⁶ Strictly speaking, it may be the case that several components of the dependency graph of $\{A_1 \dots A_{m+\delta}\}$ form a component of the dependency graph of Σ . However, it does not affect the results.

considering the components separately when dealing with nonnegative matrices, which then makes the two methods as effective as each other. At any rate, Theorem 2 gives explicit constants. Another drawback of the bound in (1) is that we do not yet have an effective method to compute or estimate (from below) the ordinary spectral radius in the first place. However, an advantage of the bound in (1) is that it works for any complex matrices, not necessarily nonnegative ones. An open problem in this direction is how our results would be extended for more general matrices.

Before closing the paper, we give an example where Theorem 2 performs better than the method in (1). For simplicity, we consider the case where the set contains only one matrix A with its powers as follows:

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad A^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}.$$

Obviously, $\rho(A) = 1$. For every n , the bounds in (1) become

$$1 \leq \rho(A) \leq \sqrt[n]{2n}.$$

Meanwhile, the bounds in Theorem 2 become

$$\sqrt[n]{1/4} \leq \rho(A) \leq \sqrt[n]{2}.$$

The ratio of the upper bound and the lower bound is respectively $\sqrt[n]{2n}$ and $\sqrt[n]{8}$. The latter converges to $\rho(A) = 1$ faster than the former.

Declaration of competing interest

The author declares that there is no competing interest.

Data availability

No data was used for the research described in the article.

References

- [1] Gian-Carlo Rota, Gilbert Strang, A note on the joint spectral radius, *Indag. Math.* 22 (4) (1960) 379–381.
- [2] Raphaël Jungers, *The Joint Spectral Radius: Theory and Applications*, vol. 385, Springer Science & Business Media, 2009.
- [3] Michael Fekete, Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten, *Math. Z.* 17 (1) (1923) 228–249.
- [4] Ingrid Daubechies, Jeffrey C. Lagarias, Sets of matrices all infinite products of which converge, *Linear Algebra Appl.* 161 (1992) 227–263.
- [5] Marc A. Berger, Yang Wang, Bounded semigroups of matrices, *Linear Algebra Appl.* 166 (1992) 21–27.
- [6] Gustaf Gripenberg, Computing the joint spectral radius, *Linear Algebra Appl.* 234 (1996) 43–60.

- [7] Victor Kozyakin, On accuracy of approximation of the spectral radius by the Gelfand formula, *Linear Algebra Appl.* 431 (11) (2009) 2134–2141.
- [8] Fabian Wirth, On the calculation of time-varying stability radii, *Int. J. Robust Nonlinear Control* 8 (12) (1998) 1043–1058.
- [9] John N. Tsitsiklis, Vincent D. Blondel, The Lyapunov exponent and joint spectral radius of pairs of matrices are hard—when not impossible—to compute and to approximate, *Math. Control Signals Syst.* 10 (1) (1997) 31–40.
- [10] Vincent D. Blondel, John N. Tsitsiklis, The boundedness of all products of a pair of matrices is undecidable, *Syst. Control Lett.* 41 (2) (2000) 135–140.
- [11] Vincent D. Blondel, Yuri Nesterov, Jacques Theys, On the accuracy of the ellipsoid norm approximation of the joint spectral radius, *Linear Algebra Appl.* 394 (2005) 91–107.