ELA

http://math.technion.ac.il/iic/ela

ON THE TRACE CHARACTERIZATION OF THE JOINT
SPECTRAL RADIUS[*]

JIANHONG XU[†]

Abstract. A characterization of the joint spectral radius, due to Chen and Zhou, relies on the limit superior of the k-th root of the dominant trace over products of matrices of length k. In this note, a sufficient condition is given such that the limit superior takes the form of a limit. This result is useful while estimating the joint spectral radius. Although it applies to a restricted class of matrices, it appears to be relevant to many realistic situations.

Key words. Joint spectral radius, Spectral radius, Trace, Primitivity, Nonnegative matrix.

AMS subject classifications. 15A18, 15A69, 65F15.

1. Introduction. The joint spectral radius has drawn much attention lately, see, for example, [1, 2, 3, 4, 5, 6, 7, 9, 13, 18] and the references therein, due to its applications in various areas such as switched systems [12], differential equations [8], coding theory [14], and wavelets [15]. For more background material, we also refer the reader to the monograph [11].

As the joint spectral radius can be regarded as a generalization of the (conventional) spectral radius, we shall denote throughout both the joint spectral radius and the spectral radius by $\rho(\cdot)$. It is well known that given a set of $n \times n$ matrices $\Sigma = \{A_1, \ldots, A_m\}$, its joint spectral radius can be characterized by

$$(1.1) \qquad \rho(\Sigma) = \lim_{k \to \infty} \max_{A \in \Pi_k} \|A\|^{1/k},$$

with Π_k being the set of products of A_i of length k whose factors are in Σ and $\|\cdot\|$ being a matrix norm; or, equivalently, by

$$(1.2) \qquad \rho(\Sigma) = \limsup_{k \to \infty} \max_{A \in \Pi_k} \rho^{1/k}(A).$$

Recently, Chen and Zhou showed in [4], also see [18], that $\rho(\Sigma)$ can be expressed alternatively as

$$(1.3) \qquad \rho(\Sigma) = \limsup_{k \to \infty} \max_{A \in \Pi_k} |\mathrm{tr}(A)|^{1/k}.$$

ELA

http://math.technion.ac.il/iic/ela

This result was then used in [4, 18] to estimate $\rho(\Sigma)$. An attractive feature of (1.3) is that it characterizes $\rho(\Sigma)$ by the trace, hence it is computationally less expensive[1] as compared to (1.1) — assuming here a general subordinate matrix norm — and (1.2).

However, the utilization of (1.3) can be complicated as it involves a limit superior, which implies that, in general, the convergence in (1.3) is attained only on some subsequence $\{k_i\}$, where $k_i \to \infty$ as $i \to \infty$. It is therefore natural for us to ask when the convergence is attained on any subsequence $\{k_i\}$, i.e. in the sense of a limit.

In this short note, we shall illustrate that the limit superior in (1.3) can be replaced with a limit when Σ is a set of nonnegative matrices which has at least one primitive element. From a computational perspective, this outcome is handy since it guarantees that with any subsequence $\{k_i\}$, the corresponding

$$\{ \max_{A \in \Pi_{k_i}} |\mathrm{tr}(A)|^{1/k_i} \}$$

provides, in general, increasingly tighter approximations to $\rho(\Sigma)$ as k_i grows.

2. Trace Characterization. Given a matrix A, the notations $A \geq 0$ and $A > 0$ are interpreted in an entrywise sense. Recall that a square matrix $A \geq 0$ is called primitive if $A^k > 0$ for some integer $k \geq 1$. In the sequel, matrices are assumed to be square. Let us begin with the following lemma:

LEMMA 2.1. ([17, p.49]) *Let $A \geq 0$ be primitive. Then*

(2.1) $$\rho(A) = \lim_{k \to \infty} \mathrm{tr}^{1/k}(A^k).$$

We comment that the assumption of primitivity in Lemma 2.1 is necessary. Consider, for instance, $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, which is not primitive; clearly, $\mathrm{tr}(A^k) = 0$ or 2, depending on whether k is odd or even, thus the limit in (2.1) does not exist. Incidentally, if $A \geq 0$ is irreducible, i.e. its underlying digraph is strongly connected, and if the diagonal entries of A are not all zero, then A must be primitive. This requirement on the diagonal entries serves as a convenient, although only sufficient, condition for determining primitivity — note that, conversely, primitivity implies irreducibility.

We shall also need a conclusion from [2] as below.

---

[1]Nevertheless, the computation of the joint spectral radius remains a tough challenge since the main difficulty comes from the cardinality of Π_k as k grows, or the construction of extremal polytope norms [9], or the size of the eigenproblem as a result of Kronecker and semidefinite liftings [2].

LEMMA 2.2. ([2, Theorem 3]) *Let* $\Sigma = \{A_1, \ldots, A_m\}$ *be a set of* $n \times n$ *nonnegative matrices. Then*

$$\rho(\Sigma) = \lim_{k \to \infty} \rho^{1/k}(A_1^{\otimes k} + \ldots + A_m^{\otimes k}),$$

where $A^{\otimes k}$ *is the* k-*th Kronecker power of* A.

For background material on the Kronecker power and, more generally, Kronecker product, see [10, 16]; a few properties of such products are used in this note without being first stated as preparatory lemmas. However, we mention here that, according to [10, Lemma 4.2.10], it follows by induction that

$$(2.2) \qquad A_1^{\otimes k} \cdots A_l^{\otimes k} = (A_1 \cdots A_l)^{\otimes k}$$

holds for any positive integers k and l. In particular, $(A^{\otimes k})^l = (A^l)^{\otimes k}$.

Continuing, we point out that applying Lemma 2.1, along with Lemma 2.2, to answer our main question here hinges on the next two simple, yet useful, observations.

LEMMA 2.3. *Suppose that* $A \geq 0$ *is primitive. Then* $A^{\otimes k}$ *is also primitive for any integer* $k \geq 1$.

Proof. By the assumption, there exists some integer $l \geq 1$ such that $A^l > 0$. Hence

$$(A^{\otimes k})^l = (A^l)^{\otimes k} > 0.$$

This completes the proof. ☐

We mention in passing that Lemma 2.3 does not extend to irreducibility of matrices. Consider, again, $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, which is indeed irreducible; but it is easy to verify that, for example, $A^{\otimes 2}$ is no longer irreducible.

LEMMA 2.4. *Suppose that* $\{A_1, \ldots, A_m\}$ *is a set of nonnegative matrices of the same size, where at least one* A_i *is primitive. Then* $A_1 + \ldots + A_m$ *is also primitive.*

Proof. Without loss of generality, we assume that A_1 is primitive. Thus there exists an integer $k \geq 1$ such that $A_1^k > 0$. We notice that

$$(A_1 + \ldots + A_m)^k \geq A_1^k > 0,$$

which yields the conclusion. ☐

We are now ready to state the following result:

THEOREM 2.5. *For a set* $\Sigma = \{A_1, \ldots, A_m\}$ *of* $n \times n$ *nonnegative matrices, with at least one* A_i *being primitive,*

$$(2.3) \qquad \rho(\Sigma) = \lim_{k \to \infty} \max_{A \in \Pi_k} \operatorname{tr}^{1/k}(A),$$

J. Xu

where Π_k is the set of products of A_i of length k whose factors are in Σ.

Proof. First of all, based on Lemmas 2.3 and 2.4, $A_1^{\otimes k} + \ldots + A_m^{\otimes k}$ is primitive for any k. Therefore, Lemma 2.1 may be exploited to formulate $\rho(A_1^{\otimes k} + \ldots + A_m^{\otimes k})$.

For any integers $k,\ l \geq 1$, we observe that, using (2.2),

$$
\begin{aligned}
\mathrm{tr}\left(A_1^{\otimes k} + \ldots + A_m^{\otimes k}\right)^l &= \mathrm{tr}\left(\sum_{1 \leq i_1,\ldots,i_l \leq m} A_{i_1}^{\otimes k} \cdots A_{i_l}^{\otimes k}\right) \\
&= \mathrm{tr}\left[\sum_{1 \leq i_1,\ldots,i_l \leq m} (A_{i_1} \cdots A_{i_l})^{\otimes k}\right] \\
&= \sum_{1 \leq i_1,\ldots,i_l \leq m} \mathrm{tr}\left(A_{i_1} \cdots A_{i_l}\right)^{\otimes k} \\
&= \sum_{1 \leq i_1,\ldots,i_l \leq m} \mathrm{tr}^k\left(A_{i_1} \cdots A_{i_l}\right),
\end{aligned}
$$

i.e.

$$
(2.4) \qquad \mathrm{tr}\left(A_1^{\otimes k} + \ldots + A_m^{\otimes k}\right)^l = \sum_{A \in \Pi_l} \mathrm{tr}^k(A),
$$

with the cardinality of Π_l being $N = m^l$. Obviously, (2.4) implies

$$
(2.5) \qquad \mathrm{tr}^{1/kl}\left(A_1^{\otimes k} + \ldots + A_m^{\otimes k}\right)^l \leq m^{1/k} \max_{A \in \Pi_l} \mathrm{tr}^{1/l}(A).
$$

Assuming $\sigma(A_1^{\otimes k} + \ldots + A_m^{\otimes k}) = \{\lambda_1, \ldots, \lambda_{n^k}\}$, where $\sigma(\cdot)$ denotes the spectrum, we arrive at

$$
\mathrm{tr}(A_1^{\otimes k} + \ldots + A_m^{\otimes k})^l = \lambda_1^l + \ldots + \lambda_{n^k}^l \leq n^k \rho^l(A_1^{\otimes k} + \ldots + A_m^{\otimes k}),
$$

which, together with (2.4), leads to

$$
(2.6) \quad \max_{A \in \Pi_l} \mathrm{tr}^{1/l}(A) \leq \mathrm{tr}^{1/kl}\left(A_1^{\otimes k} + \ldots + A_m^{\otimes k}\right)^l \leq n^{1/l}\rho^{1/k}(A_1^{\otimes k} + \ldots + A_m^{\otimes k}).
$$

It now follows from (2.5) and (2.6) that for any integers $k, l \geq 1$,

$$
(2.7) \quad \frac{\mathrm{tr}^{1/kl}\left(A_1^{\otimes k} + \ldots + A_m^{\otimes k}\right)^l}{m^{1/k}} \leq \max_{A \in \Pi_l} \mathrm{tr}^{1/l}(A) \leq n^{1/l}\rho^{1/k}(A_1^{\otimes k} + \ldots + A_m^{\otimes k}).
$$

By taking the iterated limit $\lim_{k \to \infty} \lim_{l \to \infty}$ in the above inequalities and resorting to Lemmas 2.1 and 2.2, we conclude that

$$
\lim_{k \to \infty} \lim_{l \to \infty} \max_{A \in \Pi_l} \mathrm{tr}^{1/l}(A) = \rho(\Sigma),
$$

which immediately translates into (2.3). $\square$

ELA

Concerning Theorem 2.5, we remark that in view of the previous 2×2 example, (2.3) may fail without the primitivity of A_i's — if so, the limit superior formulation is required in (2.3). However, in its proof, we see that primitivity is needed merely in the left inequality of (2.7). In other words, Theorem 2.5 still holds if its assumption is relaxed to the primitivity of $A_1^{\otimes k} + \ldots + A_m^{\otimes k}$ at any k. For example, take $\Sigma = \{A_1, A_2\}$, where

$$A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},$$

neither being primitive. It can be readily verified that when $l \geq 4$, either $S = \prod_{i=1}^{l/2}(A_1 A_2) > 0$ for any even l or $T = A_1 \cdot \prod_{i=1}^{\lfloor l/2 \rfloor}(A_2 A_1) > 0$ for any odd l, where $\lfloor \cdot \rfloor$ is the floor function. But then

$$(A_1^{\otimes k} + A_2^{\otimes k})^l = \sum_{1 \leq i_1, \ldots, i_l \leq 2} (A_{i_1} \cdots A_{i_l})^{\otimes k},$$

which implies that $(A_1^{\otimes k} + A_2^{\otimes k})^l$ is no less than $S^{\otimes k}$ or $T^{\otimes k}$, depending on whether l is even or odd. Thus $A_1^{\otimes k} + A_2^{\otimes k}$ is primitive for all k. In other words, Theorem 2.5 remains valid on the set $\Sigma = \{A_1, A_2\}$.

Our second observation from Theorem 2.5 is the following estimate of $\rho(\Sigma)$, which arises as we push $l \to \infty$ in (2.7):

COROLLARY 2.6. (see also [2, Theorem 3]) *Under the same assumption as in Theorem 2.5,*

(2.8)
$$\frac{\rho^{1/k}\left(A_1^{\otimes k} + \ldots + A_m^{\otimes k}\right)}{m^{1/k}} \leq \rho(\Sigma) \leq \rho^{1/k}(A_1^{\otimes k} + \ldots + A_m^{\otimes k})$$

for any integer $k \geq 1$.

While being free of products of A_i's, one drawback of (2.8) is the huge size of the eigenproblem being involved. When $n = 2$, for example, a case with $k = 15$ would exhaust the available memory on a typical modern PC platform.

As we mentioned earlier, the trace characterization is computationally more advantageous than that in (1.1) or (1.2). It is also known that the trace operator is invariant under cyclic shifts of the factors in a product of A_i's. However, the benefit from this invariance is limited. Consider all products of A_i's of length k. Let k_i be the number of A_i's. Then for each fixed set of $\{k_1, \ldots, k_m\}$, there are $\dfrac{k!}{k_1! \cdots k_m!}$ different arrangements of the factors, while under the cyclic-shift equivalence, this number is reduced by a factor of k to $\dfrac{1}{k} \cdot \dfrac{k!}{k_1! \cdots k_m!}$. This analysis shows that there

ELA

http://math.technion.ac.il/iic/ela

are effectively a total of m^k/k different arrangements of A_i's of length k. Clearly, the reduction is not significant as compared with m^k.

In addition, it is interesting to note that (2.1) can now be thought of as a special case of (2.3) when the set Σ consists of a single element A. Such an extension of (2.1) is parallel to that from the well-known Gelfand's formula

$$\rho(A) = \lim_{k \to \infty} \|A^k\|^{1/k}$$

to formula (1.1).

Finally, we give an example which implements (2.3) in Matlab. Again, set $\Sigma = \{A_1, A_2\}$, where

$$A_1 = \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right] \quad \text{and} \quad A_2 = \left[\begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array} \right].$$

It is known [11, p.70] that $\rho(\Sigma) = \left(\dfrac{3 + \sqrt{13}}{2} \right)^{1/4} \approx 1.3481$.
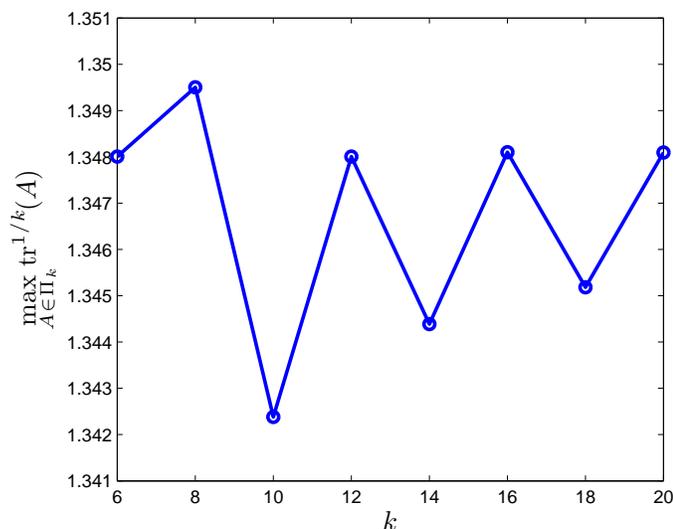

FIG. 2.1. *The curve represents* $\max_{A \in \Pi_k} \mathrm{tr}^{1/k}(A)$, *which approaches* $\rho(\Sigma)$ *as k increases. The computation is carried out in Matlab on a Dell Precision workstation with dual Xeon CPU's at 3.0 GHz and 2 GB RAM. The total running time is approximately* 250 *seconds with four pools in Matlab's Parallel Computing Toolbox.*

Consider next possible arrangements of A_i's of length k. Any such arrangement may be represented as a zero-one word in which each zero and one stand for A_1 and

A_2, respectively. All such words can be obtained by calling `de2bi(0:2^k-1,k,2)`. The result is a $2^k \times k$ array, denoted here by `u`, where each row translates to a word — we choose to generate the words at once without using a loop for the sake of speed. For the same reason, the arrangement corresponding to word `u(i,:)` are constructed by

$$\text{kron(1-u(i,:),A\_1)+kron(u(i,:),A\_2),}$$

which produces a $2 \times 2k$ array, with all the factors being arranged in this array as 2×2 blocks. Finally, the product of these factors is computed by using a loop. The numerical result on $\max_{A \in \Pi_k} \text{tr}^{1/k}(A)$ is shown in Figure 2.1, while the Matlab code for this experiment are given in Appendix A.

We mention that the above vectorized implementation runs faster than loops, but it consumes a large amount of memory. To address this issue, we may generate one `u(i,:)` at a time — the trade-off, of course, is the loss of speed.

3. Concluding Remarks. It is demonstrated here the limit superior in (1.3) can be replaced by a limit under certain circumstances. Particularly, the presence of primitivity turns out to be a sufficient condition. The lack of premitivity generally entails the limit superior formulation.

The significance of this result lies in the fact that $\max_{A \in \Pi_k} \text{tr}^{1/k}(A)$ now approaches $\rho(\Sigma)$ from both above and below. Accordingly, any subsequence of $\{k\}$ may be used to estimate $\rho(\Sigma)$ with the trace characterization. Seeing the role that nonnegative matrices play in many applications, we feel that this note is relevant to tackling problems which involve the joint spectral radius.

The trace characterization as in (2.3) can be efficiently implemented in Matlab as shown by the numerical example. Even though the applicability of this method is largely limited by the nature of the joint spectral radius, the approach we discuss here appears to be useful for solving small-scale problems concerning the joint spectral radius as well as for estimating this crucial quantity for problems of larger size by resorting to smaller k-values in (2.3).

Appendix A. Matlab Code.

```
%estimating joint spectral radius with trace characterization
clear all
a1=[0 1; 1 0];
a2=[1 1; 0 1];
k=6:2:20;
m=length(k);
jsr=zeros(m,1); %output
```

ELA

http://math.technion.ac.il/iic/ela

```
%matlabpool open 4 %parallel computing toolbox with 4 labs
for i=1:m %parfor i=1:m %with parallel computing toolbox
    trmax=0;
    n=2^k(i);
    arr=de2bi(0:n-1,k(i),2); %all zero-one words
    for j=1:n
        a=kron(1-arr(j,:),a1)+kron(arr(j,:),a2);
        u=1:2; %product
        p=a(:,u);
        for v=2:k(i)
            u=u+2;
            p=p*a(:,u);
        end
        tr=trace(p); %trace of product
        if tr>trmax
            trmax=tr;
        end
    end
    jsr(i)=trmax^(1/k(i));
end
%matlabpool close %with parallel computing toolbox
plot(k,jsr,'-bo','LineWidth',2)
```

REFERENCES

[1] M. A. Berger and Y. Wang. Bounded semigroups of matrices. *Linear Algebra Appl.*, 166:21–27, 1992.

[2] V. D. Blondel and Y. Nesterov. Computationally efficient approximations of the joint spectral radius. *SIAM J. Matrix Anal. Appl.*, 27:256–272, 2005.

[3] V. D. Blondel, J. Theys, and A. A. Vladimirov. An elementary counterexample to the finiteness conjecture. *SIAM J. Matrix Anal. Appl.*, 24:963–970, 2003.

[4] Q. Chen and X. Zhou. Characterization of joint spectral radius via trace. *Linear Algebra Appl.*, 315:175–188, 2000.

[5] I. Daubechies and J. C. Lagarias. Sets of matrices all infinite products of which converge. *Linear Algebra Appl.*, 161:227–263, 1992.

[6] I. Daubechies and J. C. Lagarias. Corrigendum/addendum to: Sets of matrices all infinite products of which converge. *Linear Algebra Appl.*, 327:69–83, 2001.

[7] L. Elsner. The generalized spectral radius theorem: An analytic-geometric proof. *Linear Algebra Appl.*, 220:151–159, 1995.

[8] N. Guglielmi and M. Zennaro. On the zero-stability of variable stepsize multistep methods: the spectral radius approach. *Numer. Math.*, 88:445–458, 2001.

[9] N. Guglielmi and M. Zennaro. An algorithm for finding extremal polytope norms of matrix families. *Linear Algebra Appl.*, 428:2265–2282, 2008.

ELA

http://math.technion.ac.il/iic/ela

[10] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis.* Cambridge University Press, Cambridge, 1991.

[11] R. Jungers. *The Joint Spectral Radius: Theory and Applications.* Springer-Verlag, Berlin, 2009.

[12] V. S. Kozyakin. Structure of extremal trajectories of discrete linear systems and the finiteness conjecture. *Automat. Remote Control*, 68:174–209, 2007.

[13] J. C. Lagarias and Y. Wang. The finiteness conjecture for the generalized spectral radius of a set of matrices. *Linear Algebra Appl.*, 214:17–42, 1995.

[14] B. Moision, A. Orlitsky, and P. H. Siegel. On codes with local joint constraints. *Linear Algebra Appl.*, 422:442–454, 2007.

[15] V. Y. Protasov. Fractal curves and wavelets. *Izv. Ross. Akad. Nauk Ser. Mat.*, 70:123–162, 2006.

[16] C. F. Van Loan. The ubiquitous Kronecker product. *J. Comput. Appl. Math.*, 123:85–100, 2000.

[17] R. S. Varga, *Matrix Iterative Analysis*, 2nd ed. Springer-Verlag, Berlin, 2000.

[18] X. Zhou, Estimates for the joint spectral radius, *Appl. Math. Comput.*, 172: 332–348, 2006.