FAST AND PRECISE APPROXIMATIONS OF THE JOINT SPECTRAL RADIUS

VINCENT D. BLONDEL AND YURII NESTEROV

Abstract. In this paper, we introduce a procedure for approximating the joint spectral radius of a finite set of matrices with arbitrary precision. Our approximation procedure is based on semidefinite liftings and can be implemented in a recursive way. For two matrices even the first step of the procedure gives an approximation, whose relative quality is at least $1/\sqrt{2}$, that is, more than 70%. The subsequent steps improve the quality but also increase the dimension of the auxiliary problem from which this approximation can be found. In an improved version of our approximation procedure we show how a relative quality of $(1/\sqrt{2})^{1/k}$ can be obtained by evaluating the spectral radius of a single matrix of dimension $n^k(n^k+1)/2$ where n is the dimension of the initial matrices. This result is computationally optimal in the sense that it provides an approximation of relative quality $1-\epsilon$ in time polynomial in $n^{1/\epsilon}$ and it is known that, unless P=NP, no such algorithm is possible that runs in time polynomial in n and $1/\epsilon$.

For the special case of matrices with non-negative entries we prove that

$$(1/2)^{1/k}\rho^{1/k}(A_1^{\otimes k}+A_2^{\otimes k}) \leq \rho(A_1,A_2) \leq \rho^{1/k}(A_1^{\otimes k}+A_2^{\otimes k})$$

where $A^{\otimes k}$ denotes the kth Kronecker power of A. An approximation of relative quality $(1/2)^{1/k}$ can therefore be obtained by computing the spectral radius of a single matrix of dimension n^k. From these inequalities it also follows that the spectral radius is given by the simple expression

$$\rho(A_1,A_2) = \lim_{k\to\infty} \|A_1^{\otimes k}+A_2^{\otimes k}\|^{1/k}$$

where it is somewhat surprising to notice that the right hand side does not directly involve any mixed products between the matrices A_1 and A_2.

This paper presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors.

Date: December 12, 2003.

1. Introduction

The spectral radius of a matrix A is an important measure of the *stability* of the discrete-time linear system

$$x_{k+1} = Ax_k, \quad x_0 \in R^n. \tag{1.1}$$

However, in some models at each step of the process (1.1) the matrix A can be chosen from a (possibly finite) *set* of matrices. The characteristic which is responsible for the stability of all possible trajectories of such a system is called the *joint spectral radius* of the set. The joint spectral radius of a set of matrices was first defined in [14] and has since then appeared in a number of different contexts; see, e.g., [15] or [7] for recent short surveys[1]. In systems theoretic terms, the joint spectral radius can be associated with the stability properties of time-varying systems in the worst case over all possible time variations, or with the stability of "asynchronous" [17] or "desynchronised" [12] systems; see also [1]. Besides systems-theoretic interpretations, the concept is pervasive in many areas of applied mathematics such as in wavelets [6], iterated function systems, fractals, numerical solutions to ordinary differential equations [9], discrete event systems [4], interpolation [18], and coding theory [13].

Despite its natural interpretation, the joint spectral radius of a set of matrices is difficult to compute. In particular, the problem of determining if the joint spectral radius of two matrices is less or equal to one is a problem that is algorithmically undecidable; see [3]. The joint spectral radius can nevertheless be approximated to any degree of accuracy by computing converging sequences of upper and lower bounds (see, e.g., [7] or [8]). Unfortunately, these bounds are expensive to compute and there are no theoretical guarantees for their rate of convergence: approximations of arbitrary degree of accuracy can be computed, but at a price that may be prohibitive. Let us say that the value ξ approximates the value ρ with *relative quality* $\mu \in [0, 1]$ (or, 100 μ %), if $\mu\, \xi \leq \rho \leq \xi$. It is known (see [16] and the note 9 on page 1260 of [5]) that, unless P=NP, there is no algorithm that can compute the joint spectral radius of two matrices with relative quality $1 - \epsilon$ in time polynomial in $1/\epsilon$. Despite this negative result it is still conceivable that for any *fixed* desired accuracy, there exists a polynomial time algorithm that computes the joint spectral radius with that accuracy. We prove in this paper that this is indeed the case. Using a certain lifting procedure (*semidefinite lifting*) we show how the joint spectral radius of m matrices can be approximated with relative quality $1/\sqrt{m}$. It appears that this procedure can be applied in a recursive way and in an improved version of the resulting approximation procedure we show how a relative quality of $(1/\sqrt{m})^{1/k}$ can be obtained by evaluating the spectral radius of a single matrix of dimension $n^k(n^k + 1)/2$. Since

$$\left(\tfrac{1}{\sqrt{m}}\right)^{1/k} \geq 1 - \tfrac{\ln m}{2k}.$$

[1]Google returns 625 entries upon entry of "joint spectral radius".

our approximation algorithm provides a relative quality $1-\epsilon$ in time polynomial in $n^{\ln m/\epsilon}$.

Contents. In Section 2 we introduce the necessary definitions and prove the lower bound on the joint spectral radius of a finite set of matrices. Further, we justify a special conic representation of the joint spectral radius. This representation is useful for the sets of matrices which leave invariant a certain cone. For such matrices we show that the usual spectral radius of the sum of these matrices is a good approximation to the joint spectral radius. In Section 3 we present the semidefinite lifting procedure, which transforms a linear operator acting in a finite-dimensional space to a linear operator acting in a space of symmetric matrices of corresponding dimension. It appears that under this transformation the joint spectral radius of a set of corresponding matrices is squared. Moreover, the lifted operators leave invariant the cone of positive semidefinite matrices. This opens a possibility to approximate the joint spectral radius of general matrices. In Section 4 we show that the semidefinite lifting can be applied recursively. We discuss the complexity issues for the case of two matrices. In Section 5 we analyze the quality of ellipsoidal bound for joint spectral radius. In Section 6 we justify another lifting procedure, the *Kronecker lifting*. We show how it works for matrices with non-negative elements. For general matrices, this procedure can be combined with semidefinite lifting, which gives more flexibility in choosing the appropriate approximation algorithm. In Section 7 we discuss the results and justify an explicit characterization for the spectral radius of a set of symmetric positive semidefinite matrices.

Notation. In this paper we denote by R^n the n-dimensional space of real column vectors. When the dimension of the space is not important, we use notation $E \equiv R^n$. For s and x in R^n we denote by $\langle s, x\rangle$ the standard inner product:

$$\langle s, x\rangle = \sum_{i=1}^{n} s^{(i)} x^{(i)}.$$

If E is equipped with a norm $\|\cdot\|$, then the *dual norm* is defined in the usual way:

$$\|s\|_* = \max_{\substack{x\in E\\ \|x\|=1}} \langle s, x\rangle.$$

For a $n\times n$ matrix A (notation: $A : E \to E$, or $A \in R^{n\times n}$) we denote by A^T its transpose. Clearly,

$$\langle s, Ax\rangle = \langle A^T s, x\rangle \quad \forall s, x \in E.$$

A *cone* in E is any subset $K \subseteq E$ such that $\lambda v \in E$ for all $\lambda \geq 0$ and $v \in K$. We say that a cone K is *proper* if it is closed, convex, has nonempty interior, and contains no line. If K is proper, then its *dual* cone

$$K^* = \{s \in E : \langle s, x\rangle \geq 0\}$$

is proper also. Note that for a matrix $A : K \to K$ we always have $A^T : K^* \to K^*$. The set of symmetric matrices, that map E to E, is denoted by $\mathbf{S} \equiv \mathbf{S}(E)$. A matrix A is positive semidefinite (which we denote by $A \succeq 0$) if $v^T A v \geq 0$ for all $v \in E$. The set of positive semidefinite matrices forms a cone in $\mathbf{S}$ which we denote by $\mathbf{S}_+$. We denote by R^n_+ the positive orthant and by Δ_n the standard simplex:

$$\Delta_n = \left\{ x \in R^n_+ : \sum_{i=1}^{n} x^{(i)} = 1 \right\}.$$

Finally, we denote by $\delta_{i,j}$ the Kronecker symbol:

$$\delta_{i,j} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

2. Definition and main properties

Let E be a finite dimensional real vector space. Consider a set of matrices $\{A_i\}_{i=1}^m$, that map E to E and let $\|\cdot\|$ be an arbitrary norm in E. Let $\{1,\dots,m\}^k$ denote the set of sequences of length k of symbols taken from the set $\{1,\dots,m\}$. To the element $\sigma \in \{1,\dots,m\}^k$ with $\sigma = \sigma_1\sigma_2\dots\sigma_k$ and $\sigma_i \in \{1,\dots,m\}$ we associate the matrix

$$S(\sigma) = A_{\sigma_k} \dots A_{\sigma_2} A_{\sigma_1}.$$

Then we can define the joint spectral radius of the set $\{A_i\}_{i=1}^m$ as follows:

$$\begin{aligned} \rho_k(A_1,\dots,A_m) &= \max_{\substack{x\in E \\ \|x\|=1}} \max_{\sigma\in\{1,\dots,m\}^k} \|S(\sigma)x\|^{1/k}, \\ &= \max_{\sigma\in\{1,\dots,m\}^k} \max_{\substack{x\in E \\ \|x\|=1}} \|S(\sigma)x\|^{1/k}, \qquad (2.1) \\ \rho(A_1,\dots,A_m) &= \limsup_{k\to\infty} \rho_k(A_1,\dots,A_m). \end{aligned}$$

Clearly, since all norms in E are equivalent, the value of the joint spectral radius does not depend on the choice of the norm. Note that for the spectral radius of a single matrix $\rho(A)$ we have the following representation:

$$\rho(A) = \max_{1\le i\le n} |\lambda_i(A)| = \lim_{k\to\infty} \max_{\substack{x\in E \\ \|x\|=1}} \|A^k x\|^{1/k}. \qquad (2.2)$$

It is well known that the value of the joint spectral radius is difficult to compute (see, e.g. [16]). Therefore a good approximation can be quite useful. First of all, consider the following *lower* bound for the joint spectral radius:

$$\tilde\rho(A_1,\dots,A_m) = \max_{\alpha\in\Delta_m} \rho\left(\sum_{i=1}^{m} \alpha^{(i)} A_i \right). \qquad (2.3)$$

The following statement was proved in [2] using the result of [12]. Here we present a direct justification.

Lemma 1. *For any set of matrices* $\{A_i\}_{i=1}^m$ *we have:*

$$\tilde\rho(A_1,\dots,A_m) \le \rho(A_1,\dots,A_m). \qquad (2.4)$$

Proof. Let us fix an arbitrary $\alpha \in \Delta_m$ and an integer $k \geq 1$. We have:

$$\|(\sum_{i=1}^{m} \alpha^{(i)} A_i)^k\| \leq \max_{\sigma \in \{1,\dots,m\}^k} \|S(\sigma)\|$$

and also

$$\limsup_{k\to\infty} \|(\sum_{i=1}^{m} \alpha^{(i)} A_i)^k\|^{1/k} \leq \limsup_{k\to\infty} \max_{\sigma \in \{1,\dots,m\}^k} \|S(\sigma)\|^{1/k}.$$

The result then follows by direct application of the relations (2.2) and (2.1).

□

Corollary 1. $\rho(A_1,\dots,A_m) \geq \frac{1}{m}\rho\left(\sum\limits_{i=1}^{m} A_i\right)$. □

In order to get an upper bound for the joint spectral radius, we introduce several useful representations of this value.

Lemma 2. *Let K be a convex cone in E with nonempty interior, and γ be a positive constant. For an integer $k \geq 1$ denote*

$$\rho_k^{[K,\gamma]}(A_1,\dots,A_m) = \max_{\substack{x\in K \\ \|x\|=\gamma}} \max_{\sigma \in \{1,\dots,m\}^k} \|S(\sigma)x\|^{1/k}.$$

Then

$$\rho(A_1,\dots,A_m) = \limsup_{k\to\infty} \rho_k^{[K,\gamma]}(A_1,\dots,A_m). \tag{2.5}$$

Proof. Since $\rho_k^{[K,\gamma]}(A_1,\dots,A_m) = \gamma^{1/k}\rho_k^{[K,1]}(A_1,\dots,A_m)$, we can restrict ourselves to the case $\gamma = 1$. It is also clear that

$$\rho_k(A_1,\dots,A_m) \geq \rho_k^{[K,1]}(A_1,\dots,A_m).$$

Therefore $\rho(A_1,\dots,A_m) \geq \rho'(A_1,\dots,A_m)$, where the latter notation stands for the right-hand side of inequality (2.5). Let us prove the converse inequality.

Consider the function

$$\|x\|_K \quad = \quad \min_{\substack{u,v\in K \\ x=u-v}} \max\{\|u\|,\|v\|\}, \quad x \in E.$$

Since $\operatorname{int} K \neq \emptyset$, this function is well defined for all $x \in E$. Moreover, it is homogeneous and convex on E, and

$$\|x\|_K \geq \min_{\substack{u,v\in E \\ x=u-v}} \max\{\|u\|,\|v\|\} \geq \tfrac{1}{2}\min_{u\in E}\{\|u\| + \|x-u\|\} = \tfrac{1}{4}\|x\|.$$

Thus, $\|\cdot\|_K$ is a well-defined norm on E.

Since the space E is finite-dimensional, there exists a constant β such that

$$\|x\|_K \leq \beta\|x\|, \quad x \in E.$$

Let us choose an arbitrary $x \in E$ with $\|x\| = 1$. Denote by $u(x)$ and $v(x)$ the points from K such that $x = u(x) - v(x)$ and $\|x\|_K = \max\{\|u(x)\|, \|v(x)\|\} \leq \beta$. Then

$$\|S(\sigma)x\| = \|S(\sigma)(u(x) - v(x))\| \leq \|S(\sigma)u(x)\| + \|S(\sigma)v(x)\|.$$

Thus, for any $x \in E$, $\|x\| = 1$ there exists a point $y \in K$ with $\|y\| = \beta$ such that

$$\|S(\sigma)x\|^{1/k} \leq 2^{1/k}\|S(\sigma)y\|^{1/k}.$$

Therefore

$$\rho_k(A_1, \ldots, A_m) \leq 2^{1/k}\rho_k^{[K,\beta]}(A_1, \ldots, A_m) = (2\beta)^{1/k}\rho_k^{[K,1]}(A_1, \ldots, A_m).$$

Hence, $\rho(A_1, \ldots, A_m) \leq \rho'(A_1, \ldots, A_m)$. □

The representation of Lemma 2 is important because of the following statement.

Theorem 1. *Assume that there exists a proper cone K such that $A_iK \subseteq K$, $i = 1, \ldots, m$. Then*

$$\frac{1}{m}\rho\left(\sum_{i=1}^{m} A_i\right) \leq \rho(A_1, \ldots, A_m) \leq \rho\left(\sum_{i=1}^{m} A_i\right). \tag{2.6}$$

Proof. Note that the lower bound in (2.6) is already established in Corollary 1. In order to prove the upper bound, let us show first that for any $x \in K$ and $s \in K^*$ we have

$$\langle s, S(\sigma)x\rangle \leq \langle s, \left(\sum_{i=1}^{m} A_i\right)^k x\rangle. \tag{2.7}$$

Indeed, using the inclusion $A_i^T s \in K^*$, $i = 1, \ldots, m$, by simple induction we obtain:

$$\begin{aligned}
\langle s, A_iS(\sigma)x\rangle &= \langle A_i^T s, S(\sigma)x\rangle \leq \langle A_i^T s, \left(\sum_{j=1}^{m} A_j\right)^k x\rangle \\
&= \langle s, A_i\left(\sum_{j=1}^{m} A_j\right)^k x\rangle \leq \langle s, \left(\sum_{j=1}^{m} A_j\right)^{k+1} x\rangle.
\end{aligned}$$

Define now the following norm:

$$\|x\|_1 = \max_{\substack{s\in K^* \\ \|s\|_*=1}} |\langle s, x\rangle|.$$

In view of inequality (2.7) for any $x \in K$ we obtain:

$$\|S(\sigma)x\|_1 = \max_{\substack{s\in K^* \\ \|s\|_*=1}} \langle s, S(\sigma)x\rangle \leq \max_{\substack{s\in K^* \\ \|s\|_*=1}} \langle s, \left(\sum_{i=1}^{m} A_i\right)^k x\rangle = \left\|\left(\sum_{i=1}^{m} A_i\right)^k x\right\|_1.$$

Thus, since the norms $\|\cdot\|$ and $\|\cdot\|_1$ are equivalent, using the representation (2.5) and property (2.2) we get

$$\begin{aligned} \rho(A_1,\dots,A_m) &\leq \limsup_{k\to\infty} \max_{\substack{x\in K\\ \|x\|=1}} \left\| \left(\sum_{i=1}^m A_i\right)^k x \right\|^{1/k} \\ &\leq \limsup_{k\to\infty} \max_{\substack{x\in E\\ \|x\|=1}} \left\| \left(\sum_{i=1}^m A_i\right)^k x \right\|^{1/k} = \rho\left(\sum_{i=1}^m A_i\right). \end{aligned}$$

□

Using the previous statement, in some simple situations we can get estimates for joint the spectral radius with relative quality $\frac{1}{m}$.

Corollary 2. *Let the matrices $A_i \in R^{n\times n}$, $i = 1,\dots,m$, have non-negative elements. Then*

$$\frac{1}{m}\rho\left(\sum_{i=1}^m A_i\right) \leq \rho(A_1,\dots,A_m) \leq \rho\left(\sum_{i=1}^m A_i\right). \tag{2.8}$$

□

However, it is more important that Theorem 1 can be applied for estimating the joint spectral radius of general matrices. We discuss this in the next section.

3. Semidefinite lifting of operators

Let A be a matrix that maps E to E and denote by $\mathbf{S}$ the space of symmetric matrices that map E to E. Consider the following object:

$$M_A(X) = AXA^T, \quad X \in \mathbf{S}.$$

Thus, $M_A(\cdot)$ is a linear operator that maps $\mathbf{S}$ to $\mathbf{S}$:

$$M_A(\cdot):\ \mathbf{S} \to \mathbf{S}.$$

We will use the notation M_A for the corresponding matrix. Note that for any positive definite matrix $X \in \mathbf{S}_+$ we have $M_A(X) \in \mathbf{S}_+$.

Having a set of matrices A_i, $i = 1,\dots,m$, we may be interested in the relation between $\rho(A_1,\dots,A_m)$ and $\rho(M_{A_1},\dots,M_{A_m})$.

Theorem 2. *For an arbitrary set of matrices $\{A_i\}_{i=1}^m$ we have*

$$\rho(M_{A_1},\dots,M_{A_m}) = \rho^2(A_1,\dots,A_m). \tag{3.1}$$

Proof. Let $n = \dim E$. For two matrices X and Y from $\mathbf{S} \equiv R^{\frac{n(n+1)}{2}}$ we introduce the Frobenius inner product and corresponding norm:

$$\langle X, Y\rangle = \sum_{i,j=1}^n X^{(i,j)}Y^{(i,j)}, \quad \|X\| = \langle X, X\rangle^{1/2}.$$

For $x, y \in E$ we choose $\|x\| = \langle x, x\rangle^{1/2}$. Consider now the values

$$\begin{aligned}
\left(\rho_k^{[\mathbf{S}_+,1]}(M_{A_1}, \ldots, M_{A_m})\right)^{2k} &= \max_{\substack{X\in\mathbf{S}_+\\ \|X\|=1}} \max_{\sigma\in\{1,\ldots,m\}^k} \|S(\sigma)XS^T(\sigma)\|^2 \\
&\geq \max_{\substack{x\in E\\ \|x\|=1}} \max_{\sigma\in\{1,\ldots,m\}^k} \|S(\sigma)\, xx^T S^T(\sigma)\|^2 \\
&= \max_{\substack{x\in E\\ \|x\|=1}} \max_{\sigma\in\{1,\ldots,m\}^k} \|S(\sigma)\, x\|^4 \\
&= (\rho_k(A_1, \ldots, A_m))^{4k}.
\end{aligned}$$

Thus, in view of Lemma 2 and definition (2.1), we get $\rho(M_{A_1}, \ldots, M_{A_m}) \geq \rho^2(A_1, \ldots, A_m)$.

On the other hand, since any $X \in \mathbf{S}_+$ with $\|X\| = 1$ can be represented as

$$X = \sum_{i=1}^n x_i x_i^T, \quad \sum_{i=1}^n \|x_i\|^4 = 1,$$

we have

$$\begin{aligned}
\max_{\substack{X\in\mathbf{S}_+\\ \|X\|=1}} \max_{\sigma\in\{1,\ldots,m\}^k} \|S(\sigma)XS^T(\sigma)\|^2 &\leq \max_{\substack{x_i\in E,\\ \sum_{i=1}^n \|x_i\|^4=1}} \max_{\sigma\in\{1,\ldots,m\}^k} \|S(\sigma) \sum_{i=1}^n x_i x_i^T S^T(\sigma)\|^2 \\
&\leq n \max_{\substack{x_i\in E,\\ \sum_{i=1}^n \|x_i\|^4=1}} \max_{\sigma\in\{1,\ldots,m\}^k} \sum_{i=1}^n \|S(\sigma)\, x_i x_i^T S^T(\sigma)\|^2 \\
&= n \max_{\substack{x\in E,\\ \|x\|=1}} \max_{\sigma\in\{1,\ldots,m\}^k} \|S(\sigma)\, xx^T S^T(\sigma)\|^2 \\
&= n\,(\rho_k(A_1, \ldots, A_m))^{4k}.
\end{aligned}$$

Thus, $\rho_k^{[\mathbf{S}_+,1]}(M_{A_1}, \ldots, M_{A_m}) \leq n^{1/(2k)} \rho^2(A_1, \ldots, A_m)$ and using again Lemma 2 and definition (2.1) we get the relation (3.1). $\square$

We can now get a simple consequence of our observations.

Theorem 3. *For the set of matrices $A_1, \ldots, A_m$ define the following linear operator:*

$$B(X) = \sum_{i=1}^m A_i X A_i^T : \quad \mathbf{S} \to \mathbf{S}.$$

Then

$$\frac{1}{\sqrt{m}}\, \rho^{1/2}(B) \leq \rho(A_1, \ldots, A_m) \leq \rho^{1/2}(B). \tag{3.2}$$

Proof. Note that the matrices M_{A_i}, $i = 1, \dots, m$, map $\mathbf{S}_+$ to $\mathbf{S}_+$, which is a proper cone. Therefore, in accordance to Theorem 1, the operator

$$B(X) \equiv \sum_{i=1}^{m} M_{A_i}(X) = \sum_{i=1}^{m} A_i X A_i^T$$

provides us with the following bounds:

$$\tfrac{1}{m}\rho(B) \leq \rho(M_{A_1}, \dots, M_{A_m}) \leq \rho(B).$$

It remains to use Theorem 2. □

Note that even for matrices with non-negative coefficients the bound (3.2) is sharper than (2.8).

4. Recursive semidefinite lifting

We now discuss the possibility of getting tighter bounds for the joint spectral radius. In order to make the presentation more transparent, we restrict ourselves to $m = 2$.

Note that for $m = 2$ the bounds (3.2) are not so bad:

$$\tfrac{1}{\sqrt{2}}\rho^{1/2}(B) \leq \rho(A_1, A_2) \leq \rho^{1/2}(B), \tag{4.1}$$

where $B(X) = A_1 X A_1^T + A_2 X A_2^T$. Indeed, $\frac{1}{\sqrt{2}} > 0.7$ and so the relative quality of our approximation is at least 70%. Note that $B(X)$ is a linear operator and so its spectral radius can be found either by standard Linear Algebra techniques, or by solving the following optimization problem:

$$\inf_{X,\tau} \left\{ \tau : \ \tau X \succeq A_1 X A_1^T + A_2 X A_2^T, \ X \succ 0 \right\}. \tag{4.2}$$

However, the main advantage of the approach presented in Section 3 consists in the possibility of applying it recursively. Indeed, denote $\rho = \rho(A_1, A_2)$ and let us introduce two sequences of linear operators $\left\{A_1^{[l]}\right\}_{l\geq 0}$ and $\left\{A_2^{[l]}\right\}_{l\geq 0}$. For $l = 0$ define

$$A_1^{[0]}(x_0) \equiv A_1 x_0, \quad A_2^{[0]}(x_0) \equiv A_2 x_0, \quad x_0 \in E_0 \equiv E.$$

For $l \geq 1$ we define these operators recursively:

$$\left.\begin{array}{rcl} A_1^{[l+1]}(x_{l+1}) & = & M_{A_1^{[l]}}(x_{l+1}) \\ & & \\ A_2^{[l+1]}(x_{l+1}) & = & M_{A_2^{[l]}}(x_{l+1}) \end{array}\right\} \quad x_{l+1} \in E_{l+1} \equiv \mathbf{S}(E_l), \quad l \geq 0.$$

From Theorem 2 we know that

$$\rho\left(A_1^{[l]}, A_2^{[l]}\right) = \rho^{2^l}.$$

On the other hand, for any $l \geq 1$ the matrices $A_i^{[l]}(\cdot)$, $i = 1, 2$, leave the cone $\mathbf{S}(E_{l-1})$ invariant. Therefore, in view of Theorem 1, we can estimate this quantity as follows:

$$\tfrac{1}{2}\rho\left(A_1^{[l]} + A_2^{[l]}\right) \leq \rho\left(A_1^{[l]}, A_2^{[l]}\right) \leq \rho\left(A_1^{[l]} + A_2^{[l]}\right).$$

Thus, after l steps, we get the following bounds:

$$\left(\tfrac{1}{2}\right)^{1/2^l} \omega_l \leq \rho \leq \omega_l,$$

where $\omega_l = \left[\rho\left(A_1^{[l]} + A_2^{[l]}\right)\right]^{1/2^l}$. Note that

$$\left(\tfrac{1}{2}\right)^{1/2^l} \geq 1 - \tfrac{\ln 2}{2^l}. \tag{4.3}$$

Thus, the improvement in the quality of our approximation is quite fast. Unfortunately, the dimension of the spaces E_l is growing much faster. Indeed,

$$n_{l+1} \equiv \dim E_{l+1} = \tfrac{1}{2} n_l (n_l + 1).$$

Therefore asymptotically we have $n_l = O\left((\tfrac{n}{2})^{2^l}\right)$, where $n = \dim E$.

Let us display the growth of the quality of approximations as compared with increase of dimension for first steps of the recursive semidefinite lifting.

Steps	Quality	$n = 2$	$n = 10$	$n = 100$
1	0.707	3	55	5050
2	0.840	6	1540	$*$
3	0.917	21	118570	$*$
4	0.957	231	$*$	$*$
5	0.978	26796	$*$	$*$

We use symbol $*$ to mark the cases for which the dimension of the auxiliary problem goes beyond the abilities of modern computers. However, note that the matrix of operator B in (4.1) is sparse. Therefore its spectral radius is computable for much higher values of n.

5. Ellipsoidal norm

Let us compare our approximation of the joint spectral radius with other possibilities. Consider the following value (see [2]):

$$\hat{\rho}(A_1, \ldots, A_m) = \left[\inf_{X,\tau}\{\tau : \ \tau X \succeq A_i X A_i^T, \ i = 1, \ldots, m, \ X \succ 0\}\right]^{1/2}, \tag{5.1}$$

which corresponds to the best ellipsoidal norm for our set of matrices. It is easy to see that

$$\rho(A_1, \ldots, A_m) \leq \hat{\rho}(A_1, \ldots, A_m).$$

Note that the spectral radius of the matrices

$$B(X) = \sum_{i=1}^{m} A_i X A_i^T : \quad \mathbf{S} \to \mathbf{S}$$

can be represented as follows:

$$\rho(B) = \inf_{X,\tau} \{\tau : \ \tau X \succeq B(X), \ X \succ 0\}. \tag{5.2}$$

Therefore, if a pair (X, τ) is feasible for the optimization problem in (5.2), then it is feasible for the optimization problem in (5.1). Therefore

$$\hat{\rho}(A_1, \dots, A_m) \leq \rho^{1/2}(B).$$

If a pair (X, τ) is feasible for the optimization problem in (5.1), then the pair $(X, m\tau)$ is feasible for the optimization problem in (5.2). Thus,

$$\rho^{1/2}(B) \leq \sqrt{m}\, \hat{\rho}(A_1, \dots, A_m).$$

Taking into account the inequality (3.2), we get the following bounds for the joint spectral radius:

$$\tfrac{1}{\sqrt{m}} \rho^{1/2}(B) \leq \rho(A_1, \dots, A_m) \leq \hat{\rho}(A_1, \dots, A_m) \leq \rho^{1/2}(B). \tag{5.3}$$

Thus, we can see that the ellipsoidal estimate $\hat{\rho}(A_1, \dots, A_m)$ also has relative quality $\frac{1}{\sqrt{m}}$. This result significantly improves the bound obtained in [2] for the ellipsoid approximation.

6. Kronecker lifting for matrices

The main drawback of recursive applications of the lifting procedure presented in Section 3 is the extremely fast growth of the dimension of the auxiliary problem. It would be interesting to find a strategy, each step of which is less aggressive. In this section we show how that can be done. Again, for the sake of simplicity, we consider the case $m = 2$.

For two matrices $A \in R^{p_1 \times q_1}$ and $B \in R^{p_2 \times q_2}$ define their *Kronecker product*, the matrix $C \in R^{p_1 p_2 \times q_1 q_2}$, as follows:

$$C = A \otimes B \equiv \begin{pmatrix} B^{(1,1)} A & \dots & B^{(1,q_2)} A \\ & \dots & \\ B^{(p_2,1)} A & \dots & B^{(p_2,q_2)} \end{pmatrix}.$$

We also define the *Kroneker power* of the matrix $A \in R^{p \times q}$:

$$A^{\otimes l} = \underbrace{A \otimes A \dots A \otimes A}_{\iota \text{ times}} \in R^{p^l \times q^l}, \quad l \geq 1.$$

There is no need for parenthesis in this expression since the Kronecker product is an associative operation. Let us prove some properties of Kronecker products. From now on we denote by $\| \cdot \|$ the standard Euclidean vector norm and the associated matrix norm:

$$\|A\| = \max_{\|x\|=1} \|Ax\|.$$

Lemma 3. *For matrices of appropriate size we have*

(1) $(A \otimes B)^T = A^T \otimes B^T$;
(2) $(A_1 \otimes B_1)(A_2 \otimes B_2) = (A_1 A_2) \otimes (B_1 B_2)$;
(3) $A^{\otimes l} B^{\otimes l} = (AB)^{\otimes l}$;
(4) $\|A^{\otimes l}\| = \|A\|^l$.

Proof. The first statement directly follows from the definition. The second statement is quite standard; a proof can be found, e.g., in [11], Section 4.2. The third statement follows from repeated application of the second. Finally, for proving the fourth statement, note that for any symmetric matrix B the eigenvalues of the matrix $B^{\otimes l}$ are formed as all possible products of eigenvalues of matrix B of length l, and so we have

$$\|A^{\otimes l}\|^2 = \max_{\|x\|=1} \langle A^{\otimes l} x, A^{\otimes l} x \rangle = \max_{\|x\|=1} \langle (A^{\otimes l})^T A^{\otimes l} x, x \rangle = \max_{\|x\|=1} \langle (A^T A)^{\otimes l} x, x \rangle = \|A\|^{2l}.$$

□

Now we can prove the main statement of this section.

Theorem 4. *For any $l \geq 1$ we have $\rho(A_1^{\otimes l}, A_2^{\otimes l}) = \rho^l(A_1, A_2)$.*

Proof. For $\sigma \in \{1,2\}^k$ define $\mathbf{S}_l(\sigma) = A_{\sigma_k}^{\otimes l} \ldots A_{\sigma_1}^{\otimes l}$. Then, in view of Lemma 3 we have

$$\|\mathbf{S}_l(\sigma)\| = \|S(\sigma)^{\otimes l}\| = \|S(\sigma)\|^l.$$

It remains to use the definition (2.1). □

We will use Theorem 4 in order to transform an estimate of the value $\rho(A_1^{\otimes l}, A_2^{\otimes l})$ into an estimate for $\rho(A_1, A_2)$. Note that the estimate of the first value can be derived from Theorem 1 provided that there exists a proper cone K, which is invariant for $A_i^{\otimes l}$, $i = 1, 2$. Consider two situations.

1. Non-negative matrices. Suppose that both matrices A_1 and A_2 have non-negative entries. Then the cone $R_+^{n^l}$ is invariant for both of them. Define $B = A_1^{\otimes l} + A_2^{\otimes l}$. Then, in view of Theorem 1 and Theorem 4 we get the following bound:

$$\left(\tfrac{1}{2}\right)^{1/l} \rho^{1/l}(B) \leq \rho(A_1, A_2) \leq \rho^{1/l}(B). \tag{6.1}$$

Relative quality of the estimate $\rho^{1/l}(B)$ satisfies the inequality

$$\left(\tfrac{1}{2}\right)^{1/l} \geq 1 - \tfrac{\ln 2}{l}. \tag{6.2}$$

In order to compute this estimate, we need to solve an eigenvalue problem for the matrix B, whose dimension is equal to n^l. Note that the estimate of the same quality is given by the value $\rho^{1/l}(\hat{B})$, where the matrix $\hat{B}$ is defined as follows

$$\hat{B}^{(i,j)} = \max\{(A_1^{\otimes l})^{(i,j)}, (A_2^{\otimes l})^{(i,j)}\}, \quad i, j = 1, \ldots, n^l.$$

2. General matrices. In general, we cannot expect that the matrices $A_1^{\otimes l}$ and $A_2^{\otimes l}$ have a joint invariant cone. However, as we have shown in Section 3, this cone can be created by semidefinite lifting. Denote

$$B(X) = M_{A_1^{\otimes l}}(X) + M_{A_2^{\otimes l}}(X).$$

Then, using Theorems 1, 2, and 4, we obtain the bound

$$\left(\tfrac{1}{2}\right)^{1/(2l)} \rho^{1/(2l)}(B) \le \rho(A_1, A_2) \le \rho^{1/(2l)}(B) \tag{6.3}$$

with relative quality

$$\left(\tfrac{1}{2}\right)^{1/(2l)} \ge 1 - \tfrac{\ln 2}{2l}. \tag{6.4}$$

In this case the dimension of the matrix B is as follows:

$$\dim B = \tfrac{1}{2} n^l (n^l + 1). \tag{6.5}$$

Thus, using this mixed strategy, we can improve the quality of approximation.

7. Discussion

The results presented in this paper open a possibility to compute approximations for the joint spectral radius with worst-case theoretical guarantees. Of course, the computational complexity of these estimates needs further examination. Indeed, the spectral structure of the matrix $A^{\otimes l}$ is very simple and this simplicity must be inherited somehow by the matrix $A_1^{\otimes l} + A_2^{\otimes l}$. It therefore appears to be an interesting problem to investigate the possibility of constructing efficient procedures for finding the spectral radius of this sum.

In order to support the above hope, let us show that a careful treatment of the structure of the above mentioned sum can lead to some explicit results (this result is also proved in [2] without the positive semi-definiteness assumption).

Theorem 5. *Let $A_i = A_i^T \succeq 0$, $i = 1, \dots, m$. Then $\rho(A_1, \dots, A_m) = \max\limits_{1 \le i \le m} \rho(A_i)$.*

Proof. Let us take into account the following facts.

1. For any set of symmetric positive semidefinite matrices $\{M_i\}_{i=1}^m$ we have

$$\begin{aligned} \max_{1\le i\le m} \rho(M_i) &= \max_{1\le i\le m} \max_{\|x\|=1} \langle M_i x, x\rangle \\ &\le \max_{\|x\|=1} \left\langle \left(\sum_{i=1}^m M_i\right) x, x\right\rangle \equiv \rho\left(\sum_{i=1}^m M_i\right) \\ &\le m \cdot \max_{1\le i\le m} \rho(M_i). \end{aligned}$$

2. Semidefinite lifting preserves the symmetry of matrices. Indeed, if the matrix $A : E \to E$ is symmetric then for any X and Y from $\mathbf{S}$ we have

$$\langle M_A(X), Y\rangle = \langle AXA, Y\rangle = \langle X, AYA\rangle = \langle M_A(Y), X\rangle.$$

Thus, the matrix M_A is also symmetric.

3. Semidefinite lifting preserves positive semidefiniteness. Indeed, if the matrix $A \in \mathbf{S}_+$, then for any X from $\mathbf{S}$ we have

$$\langle M_A(X), X\rangle = \langle AXA, X\rangle = \langle A^{1/2}XA^{1/2}, A^{1/2}XA^{1/2}\rangle = \|A^{1/2}XA^{1/2}\|^2.$$

Thus, the matrix M_A is positive semidefinite.

Applying the above observation to the recursive semidefinite lifting procedure, we conclude that

$$\rho(A_1^{[l]},\dots,A_m^{[l]}) \;\le\; \rho\left(\sum_{i=1}^m A_i^{[l]}\right) \;\le\; m \max_{1\le i\le m} \rho(A_i^{[l]}) \;=\; m[\max_{1\le i\le m} \rho(A_i)]^{2^l},$$

$$\rho(A_1^{[l]},\dots,A_m^{[l]}) \;\ge\; \tfrac{1}{m}\rho\left(\sum_{i=1}^m A_i^{[l]}\right) \;\ge\; \tfrac{1}{m} \max_{1\le i\le m} \rho(A_i^{[l]}) \;=\; \tfrac{1}{m}[\max_{1\le i\le m} \rho(A_i)]^{2^l}.$$

Since in view of Theorem 2

$$\rho(A_1^{[l]},\dots,A_m^{[l]}) = \rho^{2^l}(A_1,\dots,A_m),$$

we get the result as $l \to \infty$. □

References

[1] N. E. Barabanov. Lyapunov indicators of discrete inclusions, parts I, II and III, *Avtomatika i Telemekhanika*, **2** , 40-46, **3**, 24-29 and **5**, 17-24, 1988. Translation in *Automat. Remote Control* part I, **49**, no. 3, 152-157, part II, **49**, no. 5, 558-565, 1988.

[2] V. D. Blondel, Yu. Nesterov and J. Theys. Polynomial-time approximation algorithms for the joint spectral radius, submitted.

[3] V. D. Blondel and J. N. Tsitsiklis. The boundedness of all products of a pair of matrices is undecidable. *Systems and Control Letters*, 41:2, pp. 135-140, 2000.

[4] V. D. Blondel, S. Gaubert, and J. N. Tsitsiklis, Approximating the spectral radius of sets of matrices in the max-algebra is NP-hard, IEEE Transactions on Automatic Control, 45:9, pp. 1762-1765, 2000.

[5] V. D. Blondel and J. N. Tsitsiklis. A survey of computational complexity results in systems and control. *Automatica*, 36:9, pp. 1249-1274, 2000.

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

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

[8] G. Gripenberg, Computing the joint spectral radius, Linear Algebra Appl. 234, 43–60, 1996.

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

[10] L. Gurvits. Stability of discrete linear inclusions. *Linear Algebra Appl.*, 231, pp. 47-85, 1995.

[11] R.A. Horn and C.R. Johnson. *Topics in Matrix Analysys.* Cambrige University Press, 1991.

[12] V. S. Kozyakin. Algebraic unsolvability of problem of absolute stability of desynchronized systems. *Automation and Remote Control*, 51, pp. 754-759, 1990.

[13] B.E. Moision, A. Orlitsky, P.H. Siegel, On Codes That Avoid Specified Differences, IEEE Trans. Inform. Theory, vol. 47, no. 1, pp. 433-442, January 2001.

[14] G.-C. Rota and W. G. Strang. A note on the joint spectral radius. *Indag. Math.*, 22, pp. 379-381, 1960.

[15] G. Strang. The joint spectral radius. Commentary by Gilbert Strang on paper number 5. In "Collected works of Gian-Carlo Rota", 2001.

[16] J. N. Tsitsiklis, V. D. Blondel The Lyapunov exponent and joint spectral radius of pairs of matrices are hard – when not impossible – to compute and to approximate. Mathematics of Control, Signals, and Systems, 10, pp. 31-40, 1997.

[17] J. N. Tsitsiklis, On the stability of asynchronous iterative processes, Math. Systems Theory, **20** , 137-153, 1987.

[18] T. Yu, Bin Han and Michael Overton, Design of Hermite Subdivision Schemes aided by Spectral Radius Optimization, SIAM Journal on Scientific Computing, **25**:2, pp. 643-656, 2003.

Division of Applied Mathematics, Université catholique de Louvain, 4 avenue Georges Lemaitre, B-1348 Louvain-la-Neuve, Belgium
E-mail address: blondel@inma.ucl.ac.be

Center for Operations Research and Econometrics (CORE), Catholic University of Louvain (UCL), 34 voie du Roman Pays, 1348 Louvain-la-Neuve, Belgium
E-mail address: nesterov@core.ucl.ac.be