On accuracy of approximation of the spectral radius by the Gelfand formula

Victor Kozyakin

Institute for Information Transmission Problems, Russian Academy of Sciences, Bolshoj Karetny lane 19, Moscow 127994 GSP-4, Russia

ARTICLE INFO

Article history:
Received 20 October 2008
Accepted 5 July 2009
Available online 7 August 2009
Submitted by R.A. Brualdi

AMS classification:
15A18
15A60

PACS:
02.10.Ud
02.10.Yn

Keywords:
Infinite matrix products
Generalized spectral radius
Joint spectral radius

ABSTRACT

The famous Gelfand formula \( \rho(A) = \limsup_{n \to \infty} \frac{\|A^n\|^{1/n}}{n} \) for the spectral radius of a matrix is of great importance in various mathematical constructions. Unfortunately, the range of applicability of this formula is substantially restricted by a lack of estimates for the rate of convergence of the quantities \( \|A^n\|^{1/n} \) to \( \rho(A) \). In the paper this deficiency is made up to some extent. By using the Bochi inequalities we establish explicit computable estimates for the rate of convergence of the quantities \( \|A^n\|^{1/n} \) to \( \rho(A) \). The obtained estimates are then extended for evaluation of the joint spectral radius of matrix sets.

1. Introduction

Let \( A \) be a complex \( d \times d \) matrix and \( \| \cdot \| \) be a norm in \( \mathbb{C}^d \). As is known, the spectral radius \( \rho(A) \) of the matrix \( A \) can be expressed in terms of the norms of its powers \( \|A^n\| \) by the following Gelfand formula:
\[ \rho(A) = \lim_{n \to \infty} \|A^n\|^{1/n}, \]  
which is equivalent to the equality
\[ \rho(A) = \inf_{n \geq 1} \|A^n\|^{1/n}. \]

Nowadays, the Gelfand formula is treated as a commonly known fact and is mentioned in practically all textbooks on linear analysis without any references to the original publication, which was apparently [1].

The spectral radius of a single matrix is defined as the maximum of modulus of its eigenvalues. For matrix sets it is impossible to define the notion of the spectral radius in the same manner. In this case, the formula (1) that was taken in [2] as the basis for the definition of some quantity similar to the spectral radius.

Let \( \mathcal{A} \) be a non-empty bounded set of complex \( m \times m \) matrices. As usually, for \( n \geq 1 \) denote by \( \mathcal{A}^n \) the set of all \( n \)-products of matrices from \( \mathcal{A} \); \( \mathcal{A}^0 = I \). Given a norm \( \| \cdot \| \) in \( \mathbb{C}^d \), the limit
\[ \rho(\mathcal{A}) = \lim_{n \to \infty} \|\mathcal{A}^n\|^{1/n}, \]  
where
\[ \|\mathcal{A}^n\| = \max_{A \in \mathcal{A}} \|A\| = \max_{A_n \in \mathcal{A}} \|A_n \cdots A_2 A_1\|. \]
is called the joint spectral radius of the matrix set \( \mathcal{A} \) [2]. The limit in (2) always exists and does not depend on the norm \( \| \cdot \| \). Moreover, for any \( n \geq 1 \) the estimates \( \rho(\mathcal{A}) \leq \|\mathcal{A}^n\|^{1/n} \) hold [2], and therefore the joint spectral radius can be defined also by the following formula:
\[ \rho(\mathcal{A}) = \inf_{n \geq 1} \|\mathcal{A}^n\|^{1/n}. \]  

Since for singleton matrix sets \( \mathcal{A} = \{A\} \) the equality (2) coincides with the Gelfand formula (1) then (2) is sometimes called the generalized Gelfand formula [3]. There are also a number of different definitions [4–10] of an analog of the spectral radius for matrix sets.

In various situations it is important to know the conditions under which \( \rho(\mathcal{A}) > 0 \). As can be seen, for example, from the following inequality:
\[ \|\mathcal{A}^d\| \leq C_d \rho(\mathcal{A})\|\mathcal{A}\|^{d-1}, \]  
see [1, Theorem A], \( \rho(\mathcal{A}) = 0 \) if and only if \( \mathcal{A}^d = \{0\} \), that is if and only if the matrix set \( \mathcal{A} \) is nilpotent.

In the case of singleton matrix sets \( \mathcal{A} = \{A\} \), as is shown in a plenty of standard courses of linear analysis, the condition \( \rho(\mathcal{A}) \neq 0 \) implies the inequalities
\[ \gamma^{(1+\ln n)/n}\|\mathcal{A}^n\|^{1/n} \leq \rho(\mathcal{A}) \leq \|\mathcal{A}^n\|^{1/n} \]  
with some constant \( \gamma \in (0, 1) \). In [12, Lem. 2.3] the inequalities (5) have been extended for the case of general matrix sets:
\[ \gamma^{(1+\ln n)/n}\|\mathcal{A}^n\|^{1/n} \leq \rho(\mathcal{A}) \leq \|\mathcal{A}^n\|^{1/n}. \]  

Unfortunately, to the best of the author’s knowledge, neither exact values for \( \gamma \) nor at least effectively computable estimates for the rate of convergence of the quantities \( \|\mathcal{A}^n\|^{1/n} \) and \( \|\mathcal{A}^n\|^{1/n} \) to their limits are known. This substantially restricts the range of applicability of the formulas (1) and (2). It is not very crucial for singleton matrix sets \( \mathcal{A} = \{A\} \) since in this case the value of \( \rho(\mathcal{A}) \) can be computed by other means. However, for the case of general matrix sets the lack of estimates for the rate of convergence of the quantities \( \|\mathcal{A}^n\|^{1/n} \) to \( \rho(\mathcal{A}) \) is much more critical since in this case, as far as is known to the author, any alternative ways for evaluation of \( \rho(\mathcal{A}) \) until now are not found.

In the paper this deficiency is made up to some extent. By using the Bochi inequalities (4) we establish below explicit computable estimates for the rate of convergence of the quantities \( \|\mathcal{A}^n\|^{1/n} \) to \( \rho(\mathcal{A}) \). Apparently, these estimates are new even for the case of matrix families consisting of a single matrix.
Nevertheless, in what follows we will study only the case when $\nu$ consists of a single matrix. However, intermediate constructions from [11] contain all the information needed to find matrices and every norm $\|\cdot\|$ in $\mathbb{C}^d$, $\|A^{d}\| \leq C_d \rho(A)\|A\|^{d-1}. \tag{7}$

In [11] the value of the constant $C_d$ is given only for the case $r = 1$, that is when the matrix family $\mathcal{A}$ consists of a single matrix. However, intermediate constructions from [11] contain all the information needed to find $C_d$. This will allow to get in Section 4 an explicit expression for $C_d$.

Due to the Bochi theorem, if $\rho(A) = 0$ then $A^d = [0]$, that is the matrix set $\mathcal{A}$ is nilpotent. By (3) a converse statement is also valid: $A^d = [0]$ implies $\rho(A) = 0$. So, theoretically verification of the condition $\rho(A) = 0$ may be fulfilled in a finite number of steps: it suffices only to check that all $d$-products of matrices from $\mathcal{A}$ vanish. Of course this remark is hardly suitable in practice since even for moderate values of $d = 3, 4, r = 5, 6$ the computational burden of calculations becomes too high. Nevertheless, in what follows we will study only the case when $\rho(A) \neq 0$ or, equivalently, $A^d \neq [0]$. 

**Theorem 1.** Given $d \geq 2$, for every bounded set $\mathcal{A}$ of complex $d \times d$ matrices and every norm $\|\cdot\|$ in $\mathbb{C}^d$, $C_d^{-\sigma_d(n)/n} - n^{\nu_d(n)/n} A^n \|n/n \leq \rho(A) \leq \|A^n\|^{1/n}, \quad n = 1, 2, \ldots, \tag{8}$

where

\[ C_d = \begin{cases} 2^d - 1 & \text{for } r = 1, \\ q^{d/2} & \text{for } r > 1, \end{cases} \]

\[ \sigma_d(n) = \begin{cases} \frac{1}{2} \left( \frac{\ln n}{\ln 2} + 1 \right) \left( \frac{\ln n}{\ln 2} + 2 \right) & \text{for } d = 2, \\ \frac{(d-1)^3}{(d-2)^2} \cdot n^{\frac{\ln(d-1)}{\ln d}} & \text{for } d > 2, \end{cases} \]

\[ \nu_d(n) = \begin{cases} \frac{\ln n + 1}{d-2} \cdot n^{\frac{\ln(d-1)}{\ln d}} & \text{for } d = 2, \\ \frac{(d-1)^2}{(d-2)^2} \cdot n^{\frac{\ln(d-1)}{\ln d}} & \text{for } d > 2. \end{cases} \]

The proof of Theorem 1 is relegated to Section 3. Clearly, the statement of Theorem 1 holds also for real matrix sets.

Note that the estimates (8) are weaker than the estimates (6). It is not clear now whether it is caused by the techniques of proof of the estimates (8) or by the fact that the obtained constants $C_d, \sigma_d(n)$ and $\nu_d(n)$ are universal, that is depend neither on a matrix set nor on the choice of the norm $\|\cdot\|$. Note also that the value of the constant $C_d$ rapidly increases in $d$. That is why the estimates (8) are hardly useful in applications and sooner are of theoretical interest. Moreover, the estimates (8) are...
essentially finite-dimensional and scarcely can be extended for linear operators in infinite-dimensional spaces.

Remark, at last, that for irreducible matrix sets $\mathcal{A}$ containing more that one matrix there are valid [12, Lemma 2.3] the following, stronger than (6) or (8), estimates:

$$\gamma \frac{1}{n} \|A^n\|^{1/n} \leq \rho(A) \leq \|A^n\|^{1/n},$$

where the constant $\gamma$ can be effectively computed [13].

3. Proof of Theorem 1

The inequality $\rho(A) \leq \|A^n\|^{1/n}$ in (8) follows from (3). For $r = 1$ the value of the constant $C_d$ is found in [11]; for $r > 1$ this constant will be evaluated in Section 4.

Let us deduce some corollaries from the Bochi theorem. Firstly note that for any natural numbers $p$ and $q$ the following inequalities hold:

$$\|A^{p+q}\| \leq \|A^p\| \cdot \|A^q\|, \quad (9)$$

from which

$$\|A^p\| \leq \left(\frac{\|A^n\|}{\rho(A)}\right)^p, \quad p = 1, 2, \ldots, \quad (10)$$

Then from (4) we immediately get:

$$\|A^d\| \leq C_d \left(\rho(A)\right)^{d-1} \|A^{d-1}\|^{d-1}, \quad k = 1, 2, \ldots$$

If we denote

$$\omega_n(A) = \frac{\|A^n\|}{\left(\rho(A)\right)^n}, \quad n = 1, 2, \ldots,$$

then the latter inequalities can be rewritten in the form:

$$\omega_{d-1}(A) \leq C_d \left(\omega_{d-1}(A)\right)^{d-1}, \quad k = 1, 2, \ldots$$

Therefore, for any integer $k = 1, 2, \ldots$

$$\omega_{d-1}(A) \leq C_d \left(\omega_{d-1}(A)\right)^{d-1},$$

$$\omega_{d-2}(A) \leq C_d \left(\omega_{d-2}(A)\right)^{d-1},$$

$$\omega_{d-3}(A) \leq C_d \left(\omega_{d-3}(A)\right)^{d-1},$$

...  

By multiplying the obtained inequalities we get:

$$\omega_{d-1}(A) \leq C_d \left(\omega_{d-1}(A)\right)^{d-1}, \quad k = 1, 2, \ldots$$

Now, note that by the Bochi inequality (7)

$$\frac{1}{\rho(A)} \leq \frac{\|A\|^{d-1}}{\|A^d\|}.$$

Hence

$$1 \leq \omega_1(A) = \frac{\|A\|}{\rho(A)} \leq C_d \frac{\|A\|^d}{\|A^d\|}.$$

This allows to derive from (11) the estimate for $\omega_{d-1}(A)$ which does not contain in the right-hand part the unknown value $\rho(A)$:
\omega_p(\lambda) \leq C_d^{(k+1)(d-1)^k} \left( \frac{\|A\|}{\|A_d\|} \right)^{(d-1)^k}, \quad k = 0, 1, \ldots \tag{12}

Now, let \( n \) be an arbitrary natural number. Then there is a natural \( k \) such that \( d^k \leq n < d^{k+1} \), and consequently for \( n \) it is valid the representation

\[ n = n_k d^k + n_{k-1} d^{k-1} + \cdots + n_0, \]

where

\[ 1 \leq n_k \leq d - 1, \quad 0 \leq n_i \leq d - 1, \quad i = 1, 2, \ldots, k - 1. \tag{13} \]

Since by (9) and (10)

\[ \omega_{p+q}(\lambda) \leq \omega_p(\lambda) \cdot \omega_q(\lambda) \]

for any natural numbers \( p \) and \( q \), then

\[ \omega_n(\lambda) \leq (\omega_d(\lambda))^{n_k} \cdot (\omega_{d-1}(\lambda))^{n_{k-1}} \cdots (\omega_1(\lambda))^{n_0}. \]

By (12) from here it follows:

\[ \omega_n(\lambda) \leq C_d^{\sigma_d(n)} \left( \frac{\|A\|}{\|A_d\|} \right)^{\nu_d(n)}, \tag{14} \]

where

\[ \sigma_d(n) = \sum_{j=0}^{k} n_j \sum_{i=0}^{j} (d - 1)^i, \quad \nu_d(n) = \sum_{j=0}^{k} n_j (d - 1)^i. \tag{15} \]

Note that, by definition of the value \( \omega_n(\lambda) \), (14) is equivalent to

\[ \|A^n\| \leq C_d^{-\sigma_d(n)/n} \left( \frac{\|A\|}{\|A_d\|} \right)^{-\nu_d(n)/n} \]

and therefore to the inequality

\[ C_d^{-\sigma_d(n)/n} = \left( \frac{\|A\|}{\|A_d\|} \right)^{-\nu_d(n)/n} \]

Since this last inequality coincides with the left-hand part of (8) then to complete the proof of the theorem it remains only to get the estimates for \( \sigma_d(n) \) and \( \nu_d(n) \). By (13) and (15)

\[ \sigma_d(n) = \sum_{j=0}^{k} n_j \sum_{i=0}^{j} (d - 1)^i \leq (d - 1) \sum_{j=0}^{k} \sum_{i=0}^{j} (d - 1)^i \]

\[ = (d - 1) \sum_{j=0}^{k} (k + 1 - j) (d - 1)^j, \tag{16} \]

\[ \nu_d(n) = \sum_{j=0}^{k} n_j (d - 1)^j \leq (d - 1) \sum_{j=0}^{k} (d - 1)^j. \tag{17} \]

By definition of the number \( k \) we have \( k \leq \frac{\ln n}{\ln 2} \). Then for \( d = 2 \) from (16), (17) it follows:

\[ \sigma_2(n) \leq \frac{(k + 1)(k + 2)}{2} \leq \frac{1}{2} \left( \frac{\ln n}{\ln 2} + 1 \right) \left( \frac{\ln n}{\ln 2} + 2 \right), \]

\[ \nu_2(n) \leq k + 1 \leq \frac{\ln n}{\ln 2} + 1. \]
Represent (16), (17) for \( d > 2 \) in the form

\[
\sigma_d(n) = \sum_{j=0}^{k} n_j \sum_{i=0}^{j} (d-1)^i \leq (d-1)^{k+1} \sum_{j=0}^{k} j+1 (d-1)^j, \tag{18}
\]

\[
v_d(n) = \sum_{j=0}^{k} n_j (d-1)^j \leq (d-1)^{k+1} \sum_{j=0}^{k} \frac{1}{(d-1)^j}. \tag{19}
\]

and use the equalities

\[
\sum_{j=0}^{\infty} x^j = \frac{1}{1-x}, \quad \sum_{j=0}^{\infty} (j+1)x^j = \frac{1}{(1-x)^2}, \quad |x| < 1.
\]

By setting here \( x = \frac{1}{d-1} \), from (18), (19) we obtain:

\[
\sigma_d(n) \leq \frac{(d-1)^{k+3}}{(d-2)^2} \leq \frac{(d-1)^3}{(d-2)^2} \cdot n^{\frac{\ln(d-1)}{\ln(d)}},
\]

\[
v_d(n) \leq \frac{(d-1)^{k+2}}{(d-2)} \leq \frac{(d-1)^2}{(d-2)} \cdot n^{\frac{\ln(d-1)}{\ln(d)}}.
\]

The theorem is proved.

4. Evaluation of \( C_d \)

In [11] existence of the constant \( C_d \) is established in Theorem A, proof of which is based on Lemmas 2 and 3 cited below.

**Lemma 2** (J. Bochi). Let \( \| \cdot \|_e \) be the Euclidian norm in \( \mathbb{C}^d \). There exists \( C_0 = C_0(d) \) such that

\[
\|SA^dS^{-1}\|_e \leq C_0 \|A\|_0 \cdot \|SA^{-1}\|_e^{d-1}
\]

for every non-empty bounded set \( A \) of \( d \times d \) matrices and every matrix \( S \in GL(d) \).

Actually, in [11] under the proof of Lemma 2 it is obtained first that for every diagonal matrix \( S \in GL(d) \) the following inequality holds:

\[
\|SA^dS^{-1}\|_0 \leq d^{d-1} \|A\|_0 \cdot \|SA^{-1}\|_0^{d-1}.
\]

with the matrix norm \( \|A\|_0 = \max |a_{ij}| \).

As is known [14, Chapter 5], the following relations between the norm \( \| \cdot \|_0 \) and the Euclidean norm \( \| \cdot \|_e \) hold:

\[
\|A\|_0 \leq \|A\|_e < d\|A\|_0.
\]

from which the chain of inequalities follows:

\[
d^{-1} \cdot \|SA^dS^{-1}\|_e \leq \|SA^dS^{-1}\|_0 \leq d^{d-1} \|A\|_0 \cdot \|SA^{-1}\|_0^{d-1}
\]

\[
\leq \|SA^dS^{-1}\|_0 \leq d^{d-1} \|A\|_e \cdot \|SA^{-1}\|_e^{d-1}.
\]

that is

\[
\|SA^dS^{-1}\|_e \leq d \cdot d^{d-1} \|A\|_e \cdot \|SA^{-1}\|_e^{d-1}.
\]

The last inequality, as shown in [11] under the proof of Lemma 2, can be easily extended to the general case \( S \in GL(d) \). Therefore \( C_0 = d^d \).

Now, let us move to consideration of Lemma 3 from [11].
Lemma 3 (J. Bochi). There exists $C = C(d)$ such that, for every two norms $\| \cdot \|_1$ and $\| \cdot \|_2$ in $\mathbb{C}^d$ there is a matrix $S \in \text{GL}(d)$ such that

1. $C^{-1} \| v \|_1 \leq \| Sv \|_2 \leq \| v \|_1$ for all $v \in \mathbb{C}^d$;
2. $C^{-1} \| A \|_1 \leq \| SAS^{-1} \|_2 \leq C \| A \|_1$ for all $d \times d$ matrices $A$.

Here the second part is an immediate consequence of the first one. To evaluate the constant $C$ in the first part, first notice that whenever Lemma 3 is applied in [11], one of the two norms $\| \cdot \|_1$ or $\| \cdot \|_2$ is the Euclidean norm.

So, let us evaluate the constant $C$ under the assumption that the norm $\| \cdot \|_1$ is arbitrary while the norm $\| \cdot \|_2$ is Euclidean. This can be done by using a matrix-theoretic version of complex John's ellipsoid theorem [15]. Certainly Bochi was not aware of this technique when he wrote his paper. To be more specific, let us reproduce the argumentation from [16].

Given a norm $\| \cdot \|_1$ in $\mathbb{C}^d$, it can be represented in the form

$$\| v \|_1^2 = \sup_{\lambda \in \Lambda} \langle H_\lambda v, v \rangle, \quad v \in \mathbb{C}^d,$$

where $(\cdot, \cdot)$ is the Euclidean scalar product in $\mathbb{C}^d$ and $\{H_\lambda, \lambda \in \Lambda\}$ is a family of semidefinite matrices. But according to [15, Theorem 2.1] for any family of semidefinite matrices $\{H_\lambda, \lambda \in \Lambda\}$ there is a positive definite matrix $H$ such that

$$\langle Hv, v \rangle \leq \sup_{\lambda \in \Lambda} \langle H_\lambda v, v \rangle \leq d \langle Hv, v \rangle, \quad v \in \mathbb{C}^d.$$

Therefore

$$\langle Hv, v \rangle \leq \| v \|_1^2 \leq d \langle Hv, v \rangle, \quad v \in \mathbb{C}^d.$$

Since the matrix $H$ may be thought of as symmetric then, by setting $S = H^{1/2}$, $\| \cdot \|_2 = \sqrt{\langle \cdot, \cdot \rangle}$ and $\| Sv \|_2^2 = \langle Sv, Sv \rangle \equiv \langle H^{1/2}v, H^{1/2}v \rangle \equiv \langle Hv, v \rangle$, we obtain

$$d^{-1} \| v \|_1^2 \leq \| Sv \|_2^2 \leq \| v \|_2^2,$$

and the conclusion of Lemma 3 is valid with the constant $C = d^{1/2}$.

Now, to evaluate the value of the constant $C_d$ in Theorem A it suffices to note that due to [11] $C_d = C_d^{C_d}$ where $C_0$ and $C$ are the constants from Lemmas 2 and 3, respectively. Hence, $C_d = d^{3d/2}$.

Acknowledgements

I am grateful to Jairo Bochi for many valuable comments and especially for the idea to use John's ellipsoid theorem for estimating the constant $C$ in Lemma 3 which have allowed to improve dramatically the estimate of $C_d$.

References


