ISSN 1064–5624, Doklady Mathematics, 2009, Vol. 80, No. 1, pp. 487–491. © Pleiades Publishing, Ltd., 2009.
Original Russian Text © V.S. Kozyakin, 2009, published in Doklady Akademii Nauk, 2009, Vol. 427, No. 2, pp. 160–164.

On the Computational Aspects of the Theory of Joint Spectral Radius

V. S. Kozyakin

Presented by Academician N.A. Kuznetsov December 2, 2008

Received December 5, 2008

DOI: 10.1134/S1064562409040097

1. STATEMENT OF THE PROBLEM

Various problems of control theory [1], nonautonomous and set-valued linear dynamical systems [2], wavelet theory [3], and many other areas of mathematics lead to the necessity of analyzing the rate of growth of products of matrices from given sets. One of the characteristics describing the exponential rate of growth of matrix products is the so-called joint, or generalized, spectral radius. The problem of obtaining effective estimates for joint spectral radius is considered is this paper.

Let $\mathcal{A} = \{A_1, A_2, \ldots, A_r\}$ be a set of $d \times d$ matrices with elements from the field $\mathbb{K}$ real or complex numbers. By $\mathcal{A}^n$ we denote the set of all products of n matrices from $\mathcal{A}$. If $\mathbb{K}^d$ is endowed with some norm $\|\cdot\|$, then the limit

$$\rho(\mathcal{A}) = \overline{\lim_{n\to\infty}} \|\mathcal{A}^n\|^{\frac{1}{n}}, \qquad (1)$$

where $\|\mathcal{A}^n\| = \max\limits_{A \in \mathcal{A}^n} \|A\|$, is called the joint spectral radius of the set of matrices $\mathcal{A}$ [4]. Formula (1) is often referred to as the generalized Gelfand formula.

In addition to (1), there are many other formulas for calculating $\rho(\mathcal{A})$. Thus, in [4, 5], it was proved that $\rho(\mathcal{A}) = \inf\limits_{\|\cdot\|} \|\mathcal{A}\|$, where the infimum is over all norms on $\mathbb{K}^d$. For irreducible sets of matrices $\mathcal{A}$, the infimum in the last equality is attained. (Recall that a set of matrices $\mathcal{A}$ is said to be irreducible if matrices from $\mathcal{A}$ have no common invariant subspaces other than $\{0\}$ and $\mathbb{K}^d$.) According to [2], for irreducible sets of matrices, the number ρ coincides with $\rho(\mathcal{A})$ if and only if there exists a norm $\|\cdot\|$ (a Barabanov norm) on $\mathbb{K}^d$ for which

$$\rho\|x\| \equiv \max_i \|A_i x\|.$$

A dual result was obtained in [6], where it was shown that a number ρ coincides with $\rho(\mathcal{A})$ if and only if, for some centrally symmetric body S (a set is called a body if it contains at least one interior point), we have $\rho S = \operatorname{conv}\left(\bigcup_{i=1}^{r} A_i S\right)$, where conv(·) denotes the convex hull of a set. In [7], it was shown that $\rho(\mathcal{A}) = \overline{\lim\limits_{n\to\infty}} \max\limits_{A \in \mathcal{A}^n} (\rho(A))^{\frac{1}{n}}$, and according to [8], $\rho(\mathcal{A}) = \overline{\lim\limits_{n\to\infty}} \max\limits_{A \in \mathcal{A}^n} |\operatorname{tr}(A)|^{\frac{1}{n}}$, where tr(·) denotes matrix trace. It was mentioned in [9] that, in the definition of joint spectral radius, norms can be replaced by positive homogeneous polynomials of even degree. In [10], it was shown that, if the elements of all matrices from $\mathcal{A}$ are nonnegative, then

$$r^{-1/n}\rho^{1/n}(\Sigma_n) \le \rho(\mathcal{A}) \le \rho^{1/n}(\Sigma_n), \qquad (2)$$

where $\Sigma_n = A_1^{\otimes n} + A_2^{\otimes n} + \ldots + A_r^{\otimes n}$ and $A^{\otimes n}$ denotes the Kronecker (tensor) product of n copies of the matrix A. Theoretically, inequalities (2) make it possible to calculate the spectral radius $\rho(\mathcal{A})$ with any required accuracy. However, as n increases, the dimension of the matrix Σ_n grows so fast that, even at the moderate values $d = 3, 4$ and $r = 5, 6$, computations become impracticable. In the case where $\mathcal{A}$ consists of arbitrary matrices, a somewhat more complicated analogue of formula (2) is valid [10].

The first algorithms for calculating $\rho(\mathcal{A})$ were suggested in [1–3] and improved in [11, 12]. Unfortu-

Kharkevich Institute for Information Transmission Problems, Russian Academy of Sciences, Bol'shoi Karetnyi per. 19, Moscow, 127994 Russia; e-mail: kozyakin@iitp.ru

nately, none of these algorithms, except, possibly, that based on formula (2), give an a priori estimate for the number of steps required to achieve a given accuracy of approximation of $\rho(\mathscr{A})$. Thus, the problem of obtaining explicit estimates for an approximation of $\rho(\mathscr{A})$ remains urgent. Moreover, in many applications, of importance are estimates for the rate of convergence in the generalized Gelfand formula (1), which has a transparent interpretation in the theory of dynamical systems.

2. A PRIORI ESTIMATES

Our purpose in this section is to obtain explicit estimates for the rate of convergence of the values $\|\mathscr{A}^n\|^{\frac{1}{n}}$ to the generalized spectral radius $\rho(\mathscr{A})$.

The situation where $\rho(\mathscr{A}) = 0$ plays a somewhat special role. To determine conditions under which $\rho(\mathscr{A}) = 0$, we apply a statement from [13], according to which there exists a constant $C_d > 1$ depending only on the dimension of space such that, for any bounded set of matrices $\mathscr{A}$ and any norm $\|\cdot\|$ on $\mathbb{K}^d$, we have

$$\|\mathscr{A}^d\| \le C_d \rho(\mathscr{A}) \|\mathscr{A}\|^{d-1}. \tag{3}$$

It follows from (3) and (1) that the equality $\rho(\mathscr{A}) = 0$ is equivalent to $\mathscr{A}^d = \{0\}$, or to the nilpotency of the matrix family $\mathscr{A}$. Thus, theoretically, the condition $\rho(\mathscr{A}) = 0$ can be verified in finitely many steps: it suffices to show that all products of d matrices from $\mathscr{A}$ vanish.

If $\rho(\mathscr{A}) \ne 0$, then, for each $n \ge 1$, we have [14]

$$\gamma_*^{\frac{1+\ln n}{n}} \|\mathscr{A}^n\|^{\frac{1}{n}} \le \rho(\mathscr{A}) \le \|\mathscr{A}^n\|^{\frac{1}{n}}, \tag{4}$$

where $\gamma_* \in (0, 1)$ is some constant. Unfortunately, no exact value or at least effectively computable estimates for γ_* are given in [14]. This is remedied, to a certain extent, by the following theorem, whose proof is based on the Bochi inequalities (3).

Theorem 1. *Let $\mathscr{A} = \{A_1, A_2, \dots, A_r\}$ be any set of $d \times d$ matrices, where $d \ge 2$, with elements from the field $\mathbb{K}$, and let $\|\cdot\|$ be an arbitrary norm on $\mathbb{K}^d$. Then, for every $n = 1, 2, \dots$,*

$$C_d^{-\frac{\sigma_d(n)}{n}} \left(\frac{\|\mathscr{A}^d\|}{\|\mathscr{A}\|^d}\right)^{\frac{\nu_d(n)}{n}} \|\mathscr{A}\|^{\frac{1}{n}} \le \rho(\mathscr{A}) \le \|\mathscr{A}\|^{\frac{1}{n}}, \tag{5}$$

where $C_d = 2^d - 1$ for $r = 1$, $C_d = (2^d - 1)^{2d} d^d$ for $r > 1$, and the constants $\sigma_d(n)$ and $\nu_d(n)$ are defined by the equalities

$$\sigma_d(n) = \frac{1}{2}\log_2 2n \cdot \log_2 4n, \quad \nu_d(n) = \log_2 2n$$

for $d = 2$ and by the equalities

$$\sigma_d(n) = \frac{(d-1)^3}{(d-2)^2} \cdot n^{\frac{\ln(d-1)}{\ln d}}, \quad \nu_d(n) = \frac{(d-1)^2}{d-2} \cdot n^{\frac{\ln(d-1)}{\ln d}}$$

for $d > 2$.

Apparently, Theorem 1 is new even in the case of a set $\mathscr{A}$ consisting of one matrix. At the same time, estimates (5) are weaker than (4). It is unclear whether this is caused by the method of proof of estimates (5) or by the universality (i.e., the independence on the matrix family and on the choice of the norm$\|\cdot\|$) of the constants C_d, $\sigma_d(n)$, and $\nu_d(n)$ involved in these estimates.

We again emphasize that the amount of calculations required to estimate $\|\mathscr{A}^d\|$ rapidly increases with d; even for $d = 3, 4$ and $r = 5, 6$, they are impracticable. Therefore, estimates (5) are hardly suitable for computing the spectral radius of a matrix family in practice and may be of only theoretical interest. Moreover, estimates (5) are essentially finite-dimensional and can hardly be generalized to families of linear operators on infinite-dimensional spaces.

For irreducible matrix families $\mathscr{A}$, inequality (4) can be substantially strengthened as [14, 6] $\gamma^{1/n}\|\mathscr{A}^n\|^{1/n} \le \rho(\mathscr{A}) \le \|\mathscr{A}^n\|^{1/n}$, where $\gamma \in (0, 1)$.

Some estimates for the constant γ were obtained in [6]. In what follows, we obtain different explicit a priori estimates for the joint spectral radius of an irreducible matrix family, which we believe to be more constructive. We define the p-measure of irreducibility (in the norm $\|\cdot\|$) of a matrix family $\mathscr{A}$ as

$$\chi_p(\mathscr{A}) = \inf_{\substack{x \in \mathbb{K}^d \\ \|x\| = 1}} \sup\{t\colon S(t) \subseteq \operatorname{conv}\{\mathscr{A}_p(x) \cup \mathscr{A}_p(-x)\}\},$$

where $S(t)$ denotes the ball of radius t in the norm $\|\cdot\|$. The measure of irreducibility $\chi_p(\mathscr{A})$ was introduced and studied in [15], where it was called the quasi-controllability measure, in relation to estimating the overshooting of families of transition modes in linear control systems. The reason for our choice of the term "measure of irreducibility" for the quantity $\chi_p(\mathscr{A})$ is that, for $p \ge d - 1$, the family of matrices $\mathscr{A}$ is irreducible if and only if $\chi_p(\mathscr{A}) > 0$.

Theorem 2. *For any $n \ge 1$ and $p \ge d - 1$,*

$$(\nu_p(\mathscr{A}))^{-1/n} \|\mathscr{A}^n\|^{1/n} \le \rho(\mathscr{A}) \le \|\mathscr{A}^n\|^{1/n},$$

where $\nu_p(\mathscr{A}) = \dfrac{\max\{1, \|\mathscr{A}\|^p\}}{\chi_p(\mathscr{A})}$.

Note that the quantity $\nu_p(\mathscr{A})$ can be effectively computed directly from $\mathscr{A}$ in finitely many algebraic operations. Note also that, except in one case, Theorem 2 is valid only for matrix families containing more than one matrix. This is because the one-matrix family $\mathcal{A} = \{A\}$ can be irreducible only if the matrix A has size 2×2 and no real eigenvalues.

3. ITERATION PROCEDURES

As mentioned in the preceding section, Theorems 1 and 2 can hardly be used in practical applications even in the case of sufficiently small-size matrices. In this connection, below, we suggest two iteration procedures which not only calculate joint spectral radii with any given accuracy but also approximately construct Barabanov norms for the corresponding sets of matrices.

In what follows, we refer to a continuous function $\gamma(t, s)$, where $t, s > 0$, with the properties $\gamma(t, t) = t$ and $\min\{t, s\} < \gamma(t, s) < \max\{t, s\}$ for $t \neq s$ as an averaging function. Examples of such functions are $\gamma(t, s) = \frac{t+s}{2}$, $\gamma(t, s) = \sqrt{ts}$, and $\gamma(t, s) = \frac{2ts}{t+s}$.

Suppose that $\mathbb{K}^m$ is endowed with a norm $\|\cdot\|_0$, $e \neq 0$ is an element of $\mathbb{K}^m$ with $\|e\|_0 = 1$, and $\gamma(\cdot, \cdot)$ is an averaging function. We recursively construct a sequence of norms $\|\cdot\|_n$ by the following rules.

MR_1: Assuming that the norm $\|\cdot\|_n$ is already known, we calculate

$$\rho_n^+ = \max_{x \neq 0} \frac{\max_i \|A_i x\|_n}{\|x\|_n}, \quad \rho_n^- = \min_{x \neq 0} \frac{\max_i \|A_i x\|_n}{\|x\|_n};$$

MR_2: We set $\gamma_n = \gamma(\rho_n^-, \rho_n^+)$ and $\mu_n = \max\{\|e\|_n, \gamma_n^{-1} \max_i \|A_i e\|_n\}$ and define the new norm

$$\|x\|_{n+1} = \mu_n^{-1} \max\{\|x\|_n, \gamma_n^{-1} \max_i \|A_i x\|_n\}.$$

Theorem 3. *For any irreducible matrix set* $\mathcal{A}$ *and any averaging function* $\gamma(t, s)$, *the sequences* $\{\rho_n^\pm\}$ *constructed by the iteration procedure* MR_1, MR_2 *converge to* $\rho(\mathcal{A})$, *and the sequence of norms* $\|\cdot\|_n$ *converges uniformly on each bounded set to some Barabanov norm* $\|\cdot\|^*$ *of the matrix set* $\mathcal{A}$. *Moreover, the sequence* $\{\rho_n^-\}$ *is nondecreasing, the sequence* $\{\rho_n^+\}$ *is nonincreasing, and* $\rho_n^- \leq \rho(\mathcal{A}) \leq \rho_n^+$ *for* $n = 1, 2, \ldots$, *which provides an a posteriori estimate for the computational error of* $\rho(\mathcal{A})$.

Thus, the iteration procedure MR_1, MR_2 yields a Barabanov norm and the joint spectral radius of the matrix set $\mathcal{A}$. The procedure for constructing the numbers $\rho_n^\pm$ is similar to the method for approximate constructing joint spectral radii suggested in [11]. At the same time, apparently, the method suggested above is the first numerical method for constructing Barabanov norms.

As the current approximation to a Barabanov norm, the procedure MR_1, MR_2 uses the maximum of the values of the approximation already constructed and some auxiliary norm; for this reason, we call it the max-relaxation method. In some situations, another version of the iteration procedure for constructing Barabanov norms and calculating joint spectral radius is preferred, which is natural to call the linear relaxation method.

Again, let $\|\cdot\|_0$ be a norm on $\mathbb{K}^m$, and let $e \neq 0$ be any element of $\mathbb{K}^m$ with $\|e\|_0 = 1$. Suppose given numbers λ^- and λ^+ satisfying the conditions $0 < \lambda^- \leq \lambda^+ < 1$, which play the role of bounds for the linear relaxation parameters in what follows. We recursively construct a sequence of norms $\|\cdot\|_n$ by the following rules.

LR_1 coincides with MR_1.

LR_2: We set $\gamma_n = \max_i \|A_i e\|_n$, take an arbitrary number $\lambda_n \in [\lambda^-, \lambda^+]$, and define the new norm

$$\|x\|_{n+1} = \lambda_n \|x\|_n + (1 - \lambda_n)\gamma_n^{-1} \max_i \|A_i x\|_n.$$

Note that the norm $\|x\|_{n+1}$ is well defined for any γ_n, because, by virtue of the irreducibility of the matrix set $\mathcal{A} = \{A_1, A_2, \ldots, A_r\}$, the vectors $A_1 x, A_2 x, \ldots, A_r x$ do not vanish simultaneously for any $x \neq 0$, whence $\rho_n^- > 0$ and $\gamma_n \geq \rho_n^- \|e\|_n > 0$.

The procedure LR_1, LR_2 differs from the procedure MR_1, MR_2 only in the method for recalculating the norms $\|\cdot\|_n$ at the second step. Moreover, unlike in the procedure MR_1, MR_2, the quantities $\rho_n^\pm$ are not used to recalculate the norms $\|\cdot\|_n$; according to Theorem 4 below, they are needed only to estimate the accuracy of $\rho(\mathcal{A})$.

Theorem 4. *For any sequences of numbers* $\{\rho_n^\pm\}$ *and norms* $\|\cdot\|_n$ *defined by the iteration procedure* LR_1, LR_2, *all assertions of Theorem* 3 *are valid.*

Instead of the iteration procedure LR_1, LR_2, we could consider the formally more general procedure in which the γ_n are arbitrary numbers such that $\gamma_n \in [\rho_n^-, \rho_n^+]$ and the obtained norms are normalized by force. The corresponding rules are as follows.

LR_1' coincides with MR_1 and LR_1.

LR_2': We take any numbers $\lambda_n \in [\lambda^-, \lambda^+]$ and $\gamma_n \in [\rho_n^-, \rho_n^+]$ and define first the auxiliary norm

$$\|x\|_{n+1}' = \lambda_n \|x_n\| + (1 - \lambda_n)\gamma_n^{-1} \min_i \|A_i x\|_n$$

and then the norm

$$\|x\|_{n+1} = \frac{\|x\|_{n+1}'}{\|e\|_{n+1}'}.$$

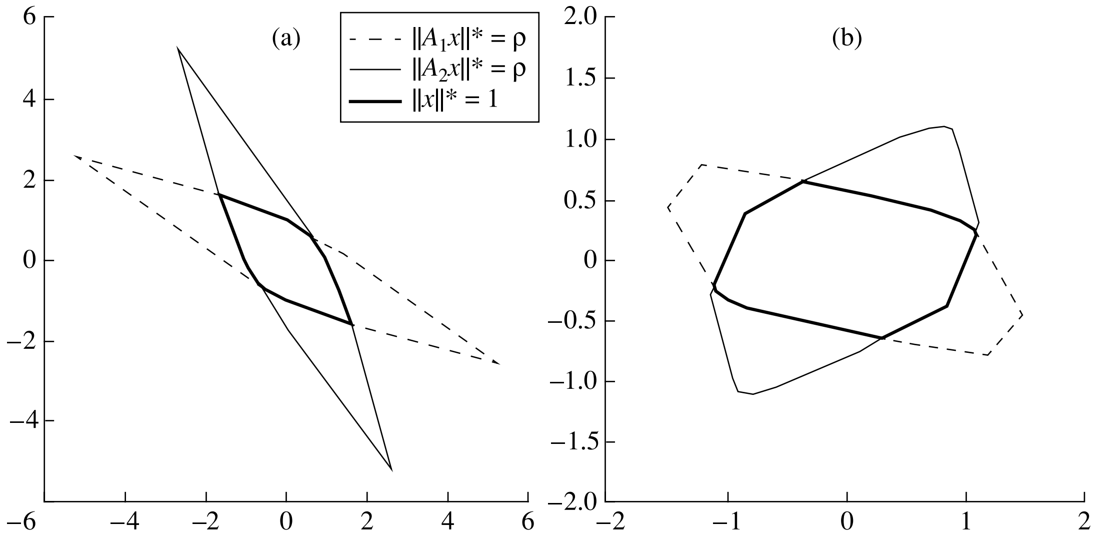

Fig. 1. Examples of calculation of Barabanov norms for a pair of two-dimensional matrices.

In fact, substituting explicit formulas for recalculating the norms $\|x\|_{n+1}$ from $\|x\|_n$ into LR_2', we obtain LR_2. Thus, the consideration of iteration procedures of the form LR_1', LR_2' does not increase generality!

4. EXAMPLES

Consider two examples of calculations by the scheme MR_1, MR_2 implemented by using the MATLAB® program. In the first example, we considered the family $\mathcal{A} = \{A_1, A_2\}$ of two-dimensional matrices $A_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $A_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Calculating $\rho(\mathcal{A})$ with absolute error 10^{-3} required 13 iterations of the algorithm MR_1, MR_2, which yielded $\rho(\mathcal{A}) = 1.617$. The approximation for the Barabanov norm $\|\cdot\|^*$ obtained after the 13th iteration of the algorithm MR_1, MR_2 is shown in Fig. 1a.

In the second example, we considered the family $\mathcal{A} = \{A_1, A_2\}$ of two-dimensional matrices $A_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $A_2 = \begin{pmatrix} \frac{4}{5} & \frac{3}{5} \\ -\frac{3}{5} & \frac{4}{5} \end{pmatrix}$. The calculation of $\rho(\mathcal{A})$ with absolute error 10^{-3} required 24 iterations of the algorithm MR_1, MR_2, which yielded $\rho(\mathcal{A}) = 1.228$. The approximation for the Barabanov norm $\|\cdot\|^*$ obtained after the 24th iteration of the algorithm MR_1, MR_2 is shown in Fig. 1b.

5. REMARKS

The question about the accuracy of the approximation of the Barabanov norms by the norms $\|\cdot\|_n$ remains open. One of the difficulties involved in obtaining the corresponding estimates is that, in the general case, Barabanov norms of a family of matrices are determined not uniquely. This is why the schemes suggested above are relaxation rather than direct. Moreover, setting $\lambda_n \equiv 0$ in the implementation of the algorithm LR_1, LR_2, we obtain a direct scheme for constructing Barabanov norms which may diverge. The question of the rate of convergence of the sequences $\{\rho_n^{\pm}\}$ to the joint spectral radius is open as well.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 06-01-00256a.

REFERENCES

1. R. K. Brayton and C. H. Tong, IEEE Trans. Circuits Syst. **27**, 1121–1130 (1980).
2. N. E. Barabanov, Avtom. Telemekh., No. 2, 40–46 (1988).
3. I. Daubechies and J. C. Lagarias, SIAM J. Math. Anal. **23**, 1031–1079 (1992).
4. G.-C. Rota and G. Strang, Indag. Math. **22**, 379–381 (1960).
5. L. Elsner, Linear Algebra Appl. **220**, 151–159 (1995).
6. V. Yu. Protasov, Fundam. Prikl. Mat. **2** (1), 205–231 (1996).
7. M. A. Berger and Y. Wang, Linear Algebra Appl. **166**, 21–27 (1992).

8. Q. Chen and X. Zhou, Linear Algebra Appl. **315**, 175–188 (2000).
9. P. A. Parrilo and A. Jadbabaie, Linear Algebra Appl. **428**, 2385–2402 (2008).
10. V. D. Blondel and Y. Nesterov, SIAM J. Matrix Anal. Appl. **27** (1), 256–272 (2005).
11. G. Gripenberg, Linear Algebra Appl. **234**, 43–60 (1996).
12. M. Maesumi, Linear Algebra Appl. **240**, 1–7 (1996).
13. J. Bochi, Linear Algebra Appl. **368**, 71–81 (2003).
14. F. Wirth, Intern. J. Robust Nonlin. Control **8**, 1043–1058 (1998).
15. V. S. Kozyakin and A. V. Pokrovskii, Dokl. Akad. Nauk **324**, 60–64 (1992).