Canad. Math. Bull. Vol. **39** (2), 1996 pp. 164–168

ON JOINT EIGENVALUES OF COMMUTING MATRICES

R. BHATIA AND L. ELSNER

ABSTRACT. A spectral radius formula for commuting tuples of operators has been proved in recent years. We obtain an analog for all the joint eigenvalues of a commuting tuple of matrices. For a single matrix this reduces to an old result of Yamamoto.

1. **Introduction, formulation of the result.** Let $T = (T_1, \ldots, T_s)$ be an s-tuple of complex $d \times d$-matrices. The *joint spectrum* $\sigma_{\mathrm{pt}}(T)$ is the set of all points $\lambda = (\lambda_1, \ldots, \lambda_s) \in C^s$ (called *joint eigenvalues*) for which there exists a nonzero vector $x \in C^d$ (called *joint eigenvector*) satisfying

$$T_j x = \lambda_j x \text{ for } j = 1, \ldots, s. \tag{1}$$

If the T_i's are commuting then $\sigma_{\mathrm{pt}}(T) \neq \emptyset$. The joint spectrum can be read off the diagonal of the common triangular form: There exists a unitary $d \times d$-matrix U such that

$$U^H T_j U = \begin{pmatrix} \lambda_1^{(j)} & \ldots & \ldots & \ldots \\ 0 & \lambda_2^{(j)} & \ldots & \ldots \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & \lambda_d^{(j)} \end{pmatrix} \text{ for } j = 1, \ldots, s. \tag{2}$$

Then

$$\sigma_{\mathrm{pt}}(T) = \{\lambda_i = (\lambda_i^{(1)}, \ldots, \lambda_i^{(s)}) : i = 1, \ldots, d\}.$$

We order the joint eigenvalues according to their norms

$$\|\lambda_1\| \geq \cdots \geq \|\lambda_d\|. \tag{3}$$

Here $\|\cdot\|$ denotes the Euclidean norm in C^r and later on also will denote the associated operator norm for matrices. We omit the reference to the dimensions.

The s-tuple T can be identified with a linear operator mapping C^d into C^{sd}. If $S = (S_1, \ldots, S_m)$ is another m-tuple of $d \times d$-matrices, we define as TS the sm-tuple of matrices, whose entries are $T_i S_j$, $i = 1, \ldots, s, j = 1, \ldots, m$, ordered lexicographically. Continuing in this way we define T^m, consisting of s^m entries, each of which is a product of m of the T_i's. Identifying again T^m with an operator mapping C^d into $C^{s^m d}$, T^m has d singular values

$$s_1(T^m) \geq s_2(T^m) \geq \cdots \geq s_d(T^m). \tag{4}$$

In this note we will prove

Research of first author supported by Sonderforschungsbereich 343 "Diskrete Modelle in der Mathematik".
Received by the editors September 11, 1995.

THEOREM 1. *For any s-tuple $T = (T_1, \ldots, T_s)$ of commuting $d \times d$-matrices*

(5) $$\lim_{m \to \infty} \left(s_j(T^m)\right)^{\frac{1}{m}} = \|\lambda_j\| \qquad j = 1, \ldots, d.$$

For $j = 1$ this has been proved in [2]; hence we know

(6) $$\|\lambda_1\| = \lim_{m \to \infty} \left(s_1(T^m)\right)^{\frac{1}{m}}.$$

We also remark that (6) has been proved in [1] for l_p-norms and in [5] for infinite-dimensional Hilbert spaces. If $s = 1$ then T^m is the usual m-th power of $T = T_1$, and the joint spectrum is the usual spectrum. For this case (5) has been proved by Yamamoto [6], who showed that for a $d \times d$ matrix T with eigenvalues λ_i ordered according to their moduli

(7) $$\lim_{m \to \infty} \left(s_j(T^m)\right)^{\frac{1}{m}} = |\lambda_j| \qquad j = 1, \ldots, d.$$

We will prove Theorem 1 in the following section.

2. **Proof of the Theorem.** It is convenient to introduce a Kronecker-type matrix product "$\tilde{\otimes}$" in the following way:

Let A and B be two (r, s) and (t, u) block matrices

$$A = (A_{ij})_{i=1,\ldots,r, j=1,\ldots,s} \qquad B = (B_{ij})_{i=1,\ldots,t, j=1,\ldots,u}$$

where the A_{ij} and B_{ij} are $d \times d$ matrices. Define

$$A_{ij}B = (A_{ij}B_{kl})_{k=1,\ldots,t, l=1,\ldots,u}$$

and the $rt \times su$-block matrix

(8) $$A \tilde{\otimes} B = \begin{pmatrix} A_{11}B & \ldots & A_{1s}B \\ \vdots & & \vdots \\ A_{r1}B & \ldots & A_{rs}B \end{pmatrix}$$

of dimension $rtd \times sud$. This product is associative. For $d = 1$ this is the usual Kronecker product, which we will denote by "$\otimes$", following the customary notation (see, *e.g.*, [4]). Except for $d = 1$, however, $A \tilde{\otimes} B$ is different from $A \otimes B$ which is an $rtd^2 \times sud^2$ matrix. So the product depends on d. However in order to avoid an overload of indices and as we keep d fixed throughout, we refrained from stressing this fact in the notation.

The main relation for $\otimes$ carries over to $\tilde{\otimes}$, namely

(9) $$(A \tilde{\otimes} B)(C \tilde{\otimes} D) = AC \tilde{\otimes} BD$$

if all the blocks in B commute with those in C, and the dimensions are fitting. For this it suffices that AC and BD can be formed. We observe that T^m, as defined in the first section, has the representation

$$T^m = T \tilde{\otimes} \cdots \tilde{\otimes} T$$

as the m-fold product of T with itself.

First we show that we can transform T to a simpler form without changing the magnitudes involved in (5). Then we prove the theorem for this simple form using (6) and (7).

Let S be a nonsingular $d \times d$ matrix,

$$\tilde{T}_i = ST_iS^{-1} \qquad i = 1, \ldots, s,$$

and

$$\tilde{T} = (\tilde{T}_1, \ldots, \tilde{T}_s).$$

Obviously the $\tilde{T}_i$'s commute too, and $\sigma_{\mathrm{pt}}(\tilde{T}) = \sigma_{\mathrm{pt}}(T)$. We show

$$(10) \qquad s_i(\tilde{T}^m) \leq \|S\| \, \|S^{-1}\| s_i(T^m) \qquad i = 1, \ldots, d,$$

which implies that the lefthand side of (5) is not changed if we replace T^m by $\tilde{T}^m$.

T^m consists of s^m blocks of $d \times d$ matrices C_i, $i = 1, \ldots, s^m$, each of which is a product of m of the T_i's. Hence the corresponding block $\tilde{C}_i$ of $\tilde{T}^m$ satisfies $\tilde{C}_i = SC_iS^{-1}$. Thus

$$(11) \qquad (\tilde{T}^m)^H \tilde{T}^m = \sum_{i=1}^{s^m} \tilde{C}_i^H \tilde{C}_i$$

$$(12) \qquad = (S^{-1})^H \Big(\sum_{i=1}^{s^m} C_i^H S^H S C_i \Big) S^{-1}$$

$$(13) \qquad \leq \|S\|^2 (S^{-1})^H (T^m)^H T^m S^{-1}$$

Here "$\leq$" is the Loewner partial ordering. Let $z \in C^d$ and $x = Sz$. The last inequality implies

$$(14) \qquad \frac{x^H (\tilde{T}^m)^H \tilde{T}^m x}{x^H x} \leq \|S\|^2 \|S^{-1}\|^2 \frac{z^H (T^m)^H T^m z}{z^H z}.$$

Using the Courant-Fischer representation of the eigenvalues $\mu_1 \geq \cdots \geq \mu_d$ of a hermitean $d \times d$ matrix B (*e.g.*, [4])

$$\mu_i = \min_{\dim V = d+1-i} \; \max_{x \in V, x \neq 0} \frac{x^H B x}{x^H x}$$

for $B = (\tilde{T}^m)^H \tilde{T}^m$ and then for $B = (T^m)^H T^m$ and taking (14) into account, (10) follows.

Another transformation of T which doesn't change the numbers $\|\lambda_i\|$ is the following:

Given a unitary $s \times s$ matrix $U = (u_{ij})$, let $W = U \otimes I_d$, where I_d is the unit matrix of dimension d, and

$$(15) \qquad \hat{T} = WT,$$

i.e.,

$$\hat{T}_i = \sum_{j=1}^{s} u_{ij} T_j \qquad i = 1, \ldots, s.$$

Then it is obvious that the joint spectrum of $\hat{T}$ is given by the vectors $\hat{\lambda}_i = U\lambda_i$, $i = 1,\dots,d$, where $\lambda_i \in \sigma_{\mathrm{pt}}(T)$. Hence $\|\hat{\lambda}_i\| = \|\lambda_i\|$, $i = 1,\dots,d$. Also by using (9) we get

$$\hat{T}^m = (WT)\tilde{\otimes}\cdots\tilde{\otimes}(WT) \tag{16}$$

$$= (W\tilde{\otimes}\cdots\tilde{\otimes}W)(T\tilde{\otimes}\cdots\tilde{\otimes}T) \tag{17}$$

$$=: W^m T^m. \tag{18}$$

Again by (9) we see that W^m defined in the last equation is a unitary mapping of $C^{s^m d}$ into itself, hence

$$s_i(\hat{T}^m) = s_i(T^m), \qquad i = 1,\dots,d.$$

Having now assembled our tools, we invoke a result in ([3], Vol. I, p. 224), by which there exists a nonsingular $d \times d$ matrix S and positive integers $s_1,\dots,s_t$ with $\sum_{i=1}^t s_i = d$, such that

$$\tilde{T}_i = ST_iS^{-1} = \operatorname{diag}(\tilde{T}_i^1,\dots,\tilde{T}_i^t) \qquad i = 1,\dots,s,$$

where

$$\tilde{T}_i^\nu = \begin{pmatrix} \tilde{\lambda}_i^\nu & \cdots & \cdots \\ 0 & \ddots & \cdots \\ 0 & 0 & \tilde{\lambda}_i^\nu \end{pmatrix} \text{ for } i = 1,\dots,s \qquad \nu = 1,\dots,t \tag{19}$$

is an $s_\nu \times s_\nu$ matrix, upper triangular with constant diagonal. Observe also that $(\tilde{T}^m)^H\tilde{T}^m$ is block diagonal with $s_\nu \times s_\nu$ blocks. This shows that we have to prove (5) only for T_i's of the form (19). Clearly then $\|\lambda_1\| = \cdots = \|\lambda_d\|$. Also by applying a suitable transformation of the form (15), we can assume that $T_2,\dots,T_d$ have zero diagonals, while the diagonal of T_1 is $\|\lambda_1\|$.

Now from

$$(T^m)^H T^m \geq (T_1^m)^H T_1^m$$

we get

$$\left(s_1(T^m)\right)^{\frac{1}{m}} \geq \left(s_i(T^m)\right)^{\frac{1}{m}} \geq \left(s_d(T_1^m)\right)^{\frac{1}{m}} \qquad i = 1,\dots,d.$$

But the leftmost term converges to $\|\lambda_1\|$ by (6), while the rightmost term converges to $\min|\lambda_i(T_1)| = \|\lambda_1\|$ by (7). Hence (5) holds for $i = 1,\dots,d$.

This finishes the proof.

REFERENCES

1. R. Bhatia and T. Bhattacharyya, *On the joint spectral radius of commuting matrices*, Studia Math. **114** (1995), 29–38.
2. M. Cho and T. Huruya, *On the joint spectral radius*, Proc. Roy. Irish Acad. Sect. A **91**(1991), 39–44.
3. F. R. Gantmacher, *The Theory of Matrices*, Chelsea, 1977.
4. M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix inequalities*, Prindle, Weber and Schmidt, Boston, 1964.

https://doi.org/10.4153/CMB-1996-020-6 Published online by Cambridge University Press

5. V. Müller and A. Soltysiak, *Spectral radius formula for commuting Hilbert space operators*, Studia Math., **103**(1992), 329–333.
6. T. Yamamoto, *On the extreme values of the roots of matrices*, J. Math. Soc. Japan **19**(1967), 173–178.

Indian Statistical Institute
Delhi centre,
7, SJS Sansanwal Marg
New Delhi 110016
India

Fakultät für Mathematik
Universität Bielefeld
Postfach 100131
D-33501 Bielefeld
Germany