MATHEMATICS

1965

SovietRxiv

Abstract

Full Text

MATHEMATICS

S. M. LOZINSKII

ON THE THEORY OF FINITE MATRICES

(Presented by Academician V. I. Smirnov on 25 I 1965)

1. Notation. C is the set of complex numbers; m, n, N, μ, ν (also with subscripts) are positive integers; $\mathbf{M}_{m\times n}$ is the set of matrices of size $m \times n$ with elements from C; $\mathbf{M}_n \stackrel{\text{def}}{=} \mathbf{M}_{n\times n}$; $\mathbf{M} \stackrel{\text{def}}{=} \bigcup_{m,n=1}^{\infty} \mathbf{M}_{m\times n}$ (i.e., $\mathbf{M}$ is the set of all finite matrices). If $A \in \mathbf{M}_{m\times n}$, then $|A|$ denotes the matrix whose elements are the moduli of the corresponding elements of A. If $A \in \mathbf{M}_n$, then A^+ denotes the matrix whose diagonal (off-diagonal) elements are the real parts (moduli) of the corresponding elements of A. If $A \in \mathbf{M}_n$, then $\rho(A)$ ($\sigma(A)$) denotes the maximum modulus (maximum real part) of the eigenvalues of the matrix A. E_n is the identity matrix of order n; $\mathfrak{O}_{m\times n}$ is the zero matrix of size $m \times n$; C^n is the set of all n-dimensional vectors with complex components. The elements of the matrix A are denoted by $a_{\mu\nu}$ or $\{A_{\mu\nu}\}$. An inequality between matrices is understood elementwise.

2. *If $A, B \in \mathbf{M}_n$, B is real and $|A| \leq B$, then $\rho(A) \leq \rho(B)$.*

Proof. If $B > \mathfrak{O}_{\times n}$, see ([1]), p. 327; for arbitrary B, by passage to the limit (the result is known).

3. Theorem 1. *If $A \in \mathbf{M}_n$, then*

$$\sigma(A) = \lim_{h\to 0+} \frac{\rho(E_n + hA) - 1}{h}. \tag{1}$$

Proof. Among the eigenvalues of the matrix A whose real part is equal to $\sigma(A)$, take the one for which the absolute value of the imaginary part is the greatest. Let $\sigma(A) + i\tau$ be this eigenvalue. Then, for all sufficiently small $h > 0$, we have $\rho(E_n + hA) = |1 + h\sigma(A) + ih\tau|$, whence (1) follows.

4. Theorem 2. *If $A, B \in \mathbf{M}_n$, B is real and $A^+ \leq B$, then $\sigma(A) \leq \sigma(B)$.*

Proof. Choose $d \geq 0$ and $h_0 > 0$ so that $|\operatorname{Im} a_{\mu\mu}| \leq d$ and $1 + h \operatorname{Re} a_{\mu\mu} > 1/2$ for $0 < h < h_0$ and $\mu = 1, \dots, n$. Then

$$|E_n + hA| \leq E_n + hB + h^2 dE_n$$

for $0 < h < h_0$, and, consequently (see p. 2),

$$\rho(E_n + hA) \leq \rho(E_n + hB + h^2 dE_n) \leq \rho(E_n + hB) + h^2 d$$

for $0 < h < h_0$. We apply Theorem 1.

5. A nonnegative function defined on $\mathbf{M}$ (on $\bigcup_{n=1}^{\infty} C^n$) is called a **matrix (vector) norm** if on each $\mathbf{M}_{m\times n}$ (on each C^n) it satisfies the three axioms of the norm of a linear complex normed space. If φ is a vector norm, then the function Φ defined on $\mathbf{M}$ by the formula

$$\Phi(A) \overset{\text{def}}{=} \sup_{\theta_n \neq x \in C^n} \frac{\varphi(Ax)}{\varphi(x)} \quad \text{for } A \in \mathbf{M}_{m\times n} \tag{2}$$

is a matrix norm (the proof is as in the case of square matrices; cf. ([2]), p. 124); we call the function Φ, defined by formula (2), the matrix norm **induced** by the vector norm φ.

6. Let $1 \le p \le +\infty$. Setting

$$\varphi(x) \overset{\text{def}}{=} \left(\sum_{k=1}^{n} |x_k|^p \right)^{1/p} \quad \text{for} \quad x = (x_1, \dots, x_n) \in C^n,$$

we obtain a vector norm. Let Φ be the matrix norm induced by the vector norm φ.

7. If φ is as in item 6; $\nu_1 + \dots + \nu_N = n$, $x = (x_{11}, \dots, x_{1\nu_1}, \dots, x_{N1}, \dots, x_{N\nu_N})$, $\xi_k = (x_{k1}, \dots, x_{k\nu_k})$ for $k = 1, \dots, N$, then

$$\varphi(x) = \varphi(\varphi(\xi_1), \dots, \varphi(\xi_N)).$$

8. Let φ and Φ be as in item 6. If $A, B \in M_{m\times n}$, $|A| \le B$, then $\Phi(A) \le \Phi(B)$. In particular, if $\mathfrak{D}_{m\times n} \le A \le B$, then $\Phi(A) \le \Phi(B)$.

The proof is clear.

9. Let $\mu_1 + \dots + \mu_N = m$; $\nu_1 + \dots + \nu_N = n$; $A \in M_{m\times n}$. Partition A into rectangular blocks A_{jk} of sizes $\mu_j \times \nu_k$ $(j, k = 1, \dots, N)$. Let φ and Φ be as in item 6. Define the matrix $B \in M_N$ by the formula

$$\{B\}_{jk} \overset{\text{def}}{=} \Phi(A_{jk}) \quad \text{for} \quad j, k = 1, \dots, N.$$

10. **Theorem 3.** With the notation of item 9 we have $\Phi(A) \le \Phi(B)$.

By item 7 this is a special case of a theorem of A. Ostrowski (([3]), p. 177, formula (III. 9)).

11. Let $\mu_1, \dots, \mu_N, \nu_1, \dots, \nu_N, \varphi, \Phi$ be as in item 9; $A \in M_{m\times n}$. Let B_p denote the matrix formed from the first $\sum_{k=1}^{p} \mu_k$ rows and the first $\sum_{k=1}^{p} \nu_k$ columns of the matrix A $(p = 1, \dots, N)$. We shall regard the matrix B_{p+1} $(p = 1, \dots, N-1)$ as consisting of four blocks: two diagonal ones (B_p and

the block $A_{p+1,p+1}$ of size $\mu_{p+1} \times \nu_{p+1}$) and two off-diagonal ones (of sizes $\mu_{p+1} \times \sum_{k=1}^{p} \nu_k$ and $(\sum_{k=1}^{p} \mu_k) \times \nu_{p+1}$). Denote the off-diagonal blocks by the symbols $C_{p+1,p}$ and $C_{p,p+1}$. Put also $\mathfrak{B}_1 \stackrel{\text{def}}{=} B_1$ and recursively define $\mathfrak{B}_2, \dots, \mathfrak{B}_N \in M_2$ by the formulas

$$\mathfrak{B}_{p+1} = \begin{bmatrix} \Phi(\mathfrak{B}_p) & \Phi(C_{p,p+1}) \\ \Phi(C_{p+1,p}) & \Phi(A_{p+1,p+1}) \end{bmatrix} \quad \text{for} \quad p = 1, \dots, N-1.$$

12. **Theorem 4.** With the notation of item 11 we have $\Phi(A) \leq \Phi(\mathfrak{B}_N)$.

Proof. By Theorem 3 and item 8 we successively obtain $\Phi(B_{p+1}) \leq \Phi(\mathfrak{B}_{p+1})$ for $p = 1, \dots, N-1$, whence (since $B_N = A$) the result follows.

13. If Φ is a matrix norm satisfying the condition $\Phi(E_n) = 1$ for $n = 1, 2, \dots$, then for any $A \in M_n$ we put

$$\gamma_\Phi(A) \stackrel{\text{def}}{=} \lim_{h \to 0+} \frac{\Phi(E_n + hA) - 1}{h} \tag{3}$$

(the limit exists; see ([4]), p. 57). The function γ_Φ, defined by formula (3) on the set of all square matrices, will be called the **logarithmic norm** corresponding to the matrix norm Φ. For properties of logarithmic norms and their applications see ([4]); let us note here only that

$$|\gamma_\Phi(A)| \leq \Phi(A)$$

(([4]), p. 58).

14. Let $\nu_1 + \dots + \nu_N = n$, $A \in M_n$, and let φ and Φ be as in item 6. Partition A into blocks A_{jk} of sizes $\nu_j \times \nu_k$ $(j, k = 1, \dots, N)$ and define the matrix $C \in M_N$ by the formula

$$\{C\}_{jk} \stackrel{\text{def}}{=} \begin{cases} \gamma_\Phi(A_{jk}), & \text{for } j = k, \\ \Phi(A_{jk}), & \text{for } j \neq k. \end{cases}$$

15. **Theorem 5.** With the notation of item 14 we have $\gamma_\Phi(A) \leq \gamma_\Phi(C)$.

Proof. Partition $E_n + hA$ into blocks of sizes $\nu_j \times \nu_k$. In this case the j-th diagonal block is $E_{\nu_j} + hA_{jj}$, and, by virtue of (3),

$$\Phi(E_{\nu_j} + hA_{jj}) = 1 + h\gamma_\Phi(A_{jj}) + o(h) \quad \text{as } h \to 0+.$$

The off-diagonal blocks have the form hA_{jk}. Using this and Theorem 3, we obtain

$$\Phi(E_n + hA) \leq \Phi(E_N + hC) + o(h) \quad \text{as } h \to 0+,$$

whence the result follows.

16. **Theorem 6.** Let $A, B \in \mathbf{M}_n$, $A^+ \leq B$; let φ and Φ be as in item 6. Then

$$\gamma_\Phi(A) \leq \gamma_\Phi(B).$$

Proof is analogous to the proof of Theorem 2, but instead of item 2, item 8 is used.

17. **Theorem 7.** Suppose that, under the conditions of item 11, we have $\mu_k = \nu_k$ for $k = 1, \dots, N$. Put $\mathfrak{C}_1 \stackrel{\text{def}}{=} B_1$, and define $\mathfrak{C}_2, \dots, \mathfrak{C}_N \in \mathbf{M}_2$ recursively by the formulas

$$\mathfrak{C}_{p+1} = \begin{bmatrix} \gamma_\Phi(\mathfrak{C}_p) & \Phi(C_{p,p+1}) \\ \Phi(C_{p+1,p}) & \gamma_\Phi(A_{p+1,p+1}) \end{bmatrix} \quad \text{for } p = 1, \dots, N-1.$$

Then

$$\gamma_\Phi(A) \leq \gamma_\Phi(\mathfrak{C}_N).$$

Proof is obtained from Theorem 5 and item 16 by the same reasoning by which Theorem 4 was obtained from Theorem 3 and item 8.

18. **Theorem 8.** Suppose that, in the notation of item 9, we have $\mu_k = \nu_k$ for $k = 1, \dots, N$. Then $\rho(A) \leq \rho(B)$.

Proof. Using block multiplication of matrices, we show that $\Phi(A^k) \leq \Phi(B^k)$ for $k = 1, 2, \dots$. We apply Theorem 6 from (5) (cf. (3), p. 185).

19. **Theorem 9.** In the notation of item 14, we have $\sigma(A) \leq \sigma(C)$.

Proof. Arguing as in the proof of Theorem 5 and applying Theorem 8, we obtain

$$\rho(E_n + hA) \leq \rho(E_N + hC + D(h)),$$

where $D(h)$ is a diagonal matrix of order N whose elements are of order $o(h)$. Since the off-diagonal elements of the matrix C are nonnegative, it follows easily (see item 2) that

$$\rho(E_N + hC + D(h)) \leq \rho(E_N + hC) + o(h).$$

Comparing the last two inequalities and using Theorem 1, we obtain the required result.

20. **Theorem 10.** Suppose that, in the notation of item 11, we have $\mu_k = \nu_k$ for $k = 1, \dots, N$. Then $\rho(A) \leq \rho(\mathfrak{B}_N)$.

Proof. By virtue of Theorem 8, we successively obtain

$$\rho(B_{p+1}) \leq \rho(\mathfrak{B}_{p+1}) \quad (p = 1, \dots, N-1).$$

But $B_N = A$.

21. **Theorem 11.** Under the conditions of Theorem 7 we have $\sigma(A) \leq \sigma(\mathfrak{C}_N)$.

Proof. By virtue of Theorem 9, we successively obtain

$$\sigma(B_{p+1}) \leq \sigma(\mathfrak{C}_{p+1}) \quad (p = 1, \dots, N-1).$$

22. The estimates given by Theorems 3 and 4, with fixed $\nu_1, \dots, \nu_N$, are incomparable. An analogous assertion is true for the pairs of Theorems 5 and 7; 8 and 10; 9 and 11.

23. Theorems 5, 8, and 9 could have been generalized (with the preservation of the proofs) to norms as general as those considered in the aforementioned (item 10) theorem of A. Ostrowski. In Theorem 3 the condition that B be a square matrix could have been dropped.

Received
17 I 1965

REFERENCES

1. F. R. Gantmacher, *The Theory of Matrices*, 1953.

2. D. K. Faddeev, V. N. Faddeeva, *Computational Methods of Linear Algebra*, 1960.

3. A. Ostrowski, *J. Math. Anal. and Appl.*, **2**, No. 2, 161 (1961).

4. S. M. Lozinskii, *Izv. Vyssh. Uchebn. Zaved. Matematika*, No. 5, 52 (1958).

5. A. Ostrowski, *Math. Zs.*, **63**, Heft 1, 2 (1955).