MATRICES OF LINEAR OPERATORS

MATRICES OF LINEAR OPERATORS [1)]

by P. M. Anselone

(Reçu le 15 octobre 1962)

In this paper we give a generalization and extension of the classical Hamilton-Cayley theorem to matrices of bounded linear operators on a Banach space. The theorem is applied to the study of the asymptotic behavior of a sequence of vectors defined by means of a composite recursion formula. In addition, the theorem is generalized to an abstract algebraic setting.

Let B be a Banach space and let $B^m = B \times \ldots \times B$ denote the product space with m factors. Elements of B^m will be denoted by row vectors

$$\vec{x} = (x_1, \ldots, x_m), \qquad x_i \in B, \tag{1}$$

or, when convenient, by column vectors. Define the norm on B^m by [2)]

$$\| \vec{x} \| = \max_i \| x_i \| . \tag{2}$$

Then B^m is a Banach space.

Let $\mathbf{T} = [T_{ij}]$ be a matrix of linear operators on B. For each $\vec{x} \in B^m$, define $\mathbf{T}\vec{x} \in B^m$ by analogy with matrix-vector multiplication:

$$\mathbf{T}\vec{x} = [T_{ij}] \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}; \tag{3}$$

more explicitly,

$$(\mathbf{T}\vec{x})_i = \sum_{j=1}^{m} T_{ij} x_j . \tag{4}$$

---

1) Sponsored by the Mathematics Research Center, United States Army, Madison, Wisconsin under Contract No.: DA-11-022-ORD-2059.

2) This particular norm is not essential for what follows. Any other equivalent norm would do, e.g., $\|\vec{x}\| = (\|x_1\|^2 + \ldots + \|x_m\|^2)^{\frac{1}{2}}$.

Thus, $\mathbf{T}$ is a linear operator on B^m. If each T_{ij} is bounded, then $\mathbf{T}$ is bounded and, by an easy argument,

$$\| \mathbf{T} \| \leqq \max_i \sum_{j=1}^{m} \| T_{ij} \| . \tag{5}$$

Recall that an operator K on a Banach space is compact (or completely continuous) if and only if it maps each bounded sequence into one with a convergent subsequence. A compact operator is necessarily bounded. It is not difficult to prove that the operator $\mathbf{K} = [K_{ij}]$ on B^m is compact if and only if each of the operators K_{ij} on B is compact.

Operators on B of the form $T = aI + K$, where a is a scalar, I is the identity operator, and K is compact are important in both theory and applications, e.g., in the study of Fredholm integral equations of the second kind. The following theorem concerns matrices of such operators.

Theorem 1. Let $\mathbf{T} = [a_{ij}I + K_{ij}]$, where the K_{ij} are compact. Let $P(\lambda)$ be the characteristic polynomial of the scalar matrix $[a_{ij}]$. Then $P(\mathbf{T})$ is compact.

Proof. Note that $\mathbf{T} = \mathbf{A} + \mathbf{K}$, where $\mathbf{A} = [a_{ij}I]$ and $\mathbf{K} = [K_{ij}]$. Then

$$P(\mathbf{T}) = P(\mathbf{A}) + Q(\mathbf{A}, \mathbf{K}),$$

where Q is a polynomial in $\mathbf{A}$ and $\mathbf{K}$, with a factor $\mathbf{K}$ in every term. Since the product of a bounded operator and a compact operator is compact, and since a sum of compact operators is compact, $Q(\mathbf{A}, \mathbf{K})$ is compact. By the Hamilton-Cayley theorem, $P([a_{ij}]) = 0$. Since the correspondence $[a_{ij}] \leftrightarrow [a_{ij}I] = \mathbf{A}$ is an algebraic isomorphism, $P(\mathbf{A}) = 0$. Therefore, $P(\mathbf{T}) = Q(\mathbf{A}, \mathbf{K})$ and, hence, $P(\mathbf{T})$ is compact.

Let $\mathbf{T}$ be as in Theorem 1. The fact that a polynomial in $\mathbf{T}$ is compact implies that $\mathbf{T}$ has a number of properties which generalize those of compact operators (cf. [1], ch. 5). We mention several of these properties. The spectrum σ of $\mathbf{T}$ is countable. The only possible limit points of σ are zeros of the characteristic polynomial $P(\lambda)$. Fix $\lambda \in \sigma$ such that $P(\lambda) \neq 0$. Then λ is an *eigenvalue* of $\mathbf{T}$. The *generalized eigenmanifolds*

$$M_\lambda^k = \{\vec{x} \in B^m : (\mathbf{T} - \lambda \mathbf{I})^k \vec{x} = 0\}, \qquad k = 0, 1, \ldots, \tag{6}$$

are finite dimensional. The ranges

$$N_\lambda^k = \{(\mathbf{T}-\lambda\mathbf{I})^k \vec{x} : \vec{x} \in B^m\}, \qquad k = 0, 1, \ldots, \tag{7}$$

are closed and have finite deficiency (codimension). There is a positive integer $\nu = \nu(\lambda)$, called the *index* of λ, such that

$$\{0\} = M_\lambda^0 \underset{\neq}{\subset} \ldots \underset{\neq}{\subset} M_\lambda^\nu = M_\lambda^{\nu+1} = \ldots, \tag{8}$$

$$B^m = N_\lambda^0 \underset{\neq}{\supset} \ldots \underset{\neq}{\supset} N_\lambda^\nu = N_\lambda^{\nu+1} = \ldots, \tag{9}$$

$$\mathbf{T} M_\lambda^\nu \subset M_\lambda^\nu, \qquad \mathbf{T} N_\lambda^\nu \subset N_\lambda^\nu, \tag{10}$$

$$B^m = M_\lambda^\nu \oplus N_\lambda^\nu. \tag{11}$$

Thus, each $\vec{x} \in B^m$ has a unique representation of the form

$$\vec{x} = \vec{u} + \vec{v}, \qquad \vec{u} \in M_\lambda^\nu, \ \vec{v} \in N_\lambda^\nu. \tag{12}$$

The restrictions of $\mathbf{T}$ to M_λ^ν and N_λ^ν have the spectra $\{\lambda\}$ and $\sigma - \{\lambda\}$, respectively. The manifolds M_λ^ν and N_λ^ν are the *spectral subspaces* associated with the subsets $\{\lambda\}$ and $\sigma - \{\lambda\}$ of σ.

Next we give an application of Theorem 1. Consider a composite recursion formula in B,

$$x_n = \sum_{j=1}^{m} T_j x_{n-j}, \qquad n = m, m+1, \ldots, \tag{13}$$

where $x_0, \ldots, x_{m-1}$, are arbitrary elements in B and the T_j, $j = 1, \ldots, m$, are bounded linear operators on B. Clearly, (13) determines x_n, $n \geqq m$ inductively in terms of $x_0, \ldots, x_{m-1}$. It is desired to study the asymptotic behavior of x_n as $n \to \infty$.

For this purpose, we let

$$\vec{x}_n = (x_n, x_{n+1}, \ldots, x_{n+m-1}) \in B^m, \qquad n = 0, 1, \ldots, \tag{14}$$

and define the bounded linear operator $\mathbf{T}$ on B^m such that

$$\mathbf{T}(x_0, \ldots, x_{m-1}) = (x_1, \ldots, x_m), \quad x_m = \sum_{j=1}^{m} T_j x_{m-j}. \tag{15}$$

Then $\vec{x}_{n+1} = T\vec{x}_n$ and, hence,

$$\vec{x}_n = T^n \vec{x}_0 , \qquad n = 1, 2, \dots . \tag{16}$$

The operator **T** has the matrix representation

$$\mathbf{T} = \begin{bmatrix} 0 & I & 0 & \dots & 0 \\ 0 & 0 & I & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & I \\ T_m & T_{m-1} & T_{m-2} & \dots & T_1 \end{bmatrix} \tag{17}$$

The method used above to replace a composite recursion formula by a simple one serves a similar purpose in the theory of multiple Markov chains (cf. [2], pp. 185-186). It is used also to replace an ordinary differential equation of m^{th} order by a system of first order equations (cf. [3], p. 82).

By (2) and (14),

$$\| \vec{x}_n \| = \max (\| x_n \| , \dots , \| x_{n+m-1} \|) , \qquad n = 0, 1, \dots . \tag{18}$$

This equation can be used to derive asymptotic results for x_n as $n \to \infty$ from corresponding results for $\vec{x}_n$. So let us consider $\vec{x}_n$.

The asymptotic behavior of $\vec{x}_n = \mathbf{T}^n\vec{x}_0$ as $n \to \infty$ is determined to a large extent by the spectral properties of **T**. This is partly because the spectral radius of **T**,

$$R_\sigma = \max \{ | \lambda | : \lambda \in \sigma \} , \tag{19}$$

satisfies the equation

$$\lim_{n \to \infty} \| \mathbf{T}^n \|^{1/n} = R_\sigma . \tag{20}$$

It follows easily from (20) that

$$\limsup_{n \to \infty} \| \mathbf{T}^n \vec{x}_0 \|^{1/n} \leqq R_\sigma . \tag{21}$$

An analogous inequality, a little more difficult to prove, is

$$\liminf_{n\to\infty} \| \mathbf{T}^n \vec{x}_0 \|^{1/n} \geqq r_\sigma \quad \text{if} \quad \vec{x}_0 \neq 0\,, \tag{22}$$

where

$$r_\sigma = \min \{ | \lambda | : \lambda \in \sigma \}\,. \tag{23}$$

Suppose now that the operators T_j in (13) are of the form $T_j = a_j\, I + K_j$, where the K_j are compact. Then, by (17), $\mathbf{T} = [a_{ij}\, I + K_{ij}]$, where the K_{ij} are compact and

$$[a_{ij}] = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 1 \\ a_m & a_{m-1} & a_{m-2} & \dots & a_1 \end{bmatrix}\,. \tag{24}$$

The characteristic polynomial of $[a_{ij}]$ is (cf. [3], pp. 88-89)

$$P(\lambda) = \lambda^m - \sum_{j=1}^{m} a_j\, \lambda^{m-j}\,. \tag{25}$$

Therefore, by Theorem 1, the operator

$$P(\mathbf{T}) = \mathbf{T}^m - \sum_{j=1}^{m} a_j\, \mathbf{T}^{m-j}\,, \tag{26}$$

is compact.

The fact that a polynomial in $\mathbf{T}$ is compact simplifies greatly the study of $\vec{x}_n = \mathbf{T}^n \vec{x}_0$ as $n \to \infty$. We consider in some detail the case with $P(\lambda) \neq 0$ for $|\lambda| = R_\sigma$. There is just a finite number $d \geqq 1$ of eigenvalues λ_k, $k = 1, \dots, d$, such that $|\lambda_k| = R_\sigma$. Let $\nu_k = \nu(\lambda_k)$ and

$$M = M_{\lambda_1}^{\nu_1} \oplus \dots \oplus M_{\lambda_d}^{\nu_d}\,, \tag{27}$$

$$N = \bigcap_{k=1}^{d} N_{\lambda_k}^{\nu_k} = \{ (\mathbf{T} - \lambda_1\, \mathbf{I})^{\nu_1} \dots (\mathbf{T} - \lambda_d\, \mathbf{I})^{\nu_d} \vec{x} : \vec{x} \in B^m \}\,. \tag{28}$$

Then

$$\mathbf{T}M \subset M , \qquad \mathbf{T}N \subset N , \tag{29}$$

$$B^m = M \oplus N . \tag{30}$$

The restrictions of $\mathbf{T}$ to M and N have the spectra $\sigma_1 = \{\lambda_i : i = 1, \dots , d\}$ and $\sigma_2 = \sigma - \sigma_1$, respectively; M and N are the spectral subspaces associated with σ_1 and σ_2.

Let

$$\vec{x}_0 = \vec{u}_0 + \vec{v}_0 , \qquad \vec{u}_0 \in M , \qquad \vec{v}_0 \in N . \tag{31}$$

Then

$$\vec{x}_n = \mathbf{T}^n \vec{x}_0 = \mathbf{T}^n \vec{u}_0 + \mathbf{T}^n \vec{v}_0 , \begin{cases} \mathbf{T}^n \vec{u}_0 \in M , \\ \mathbf{T}^n \vec{v}_0 \in N . \end{cases} \tag{32}$$

By (21) and (22), appropriately specialized,

$$\lim_{n \to \infty} \| T^n \vec{u}_0 \|^{1/n} = R_\sigma \quad \text{if} \quad \vec{u}_0 \neq 0 , \tag{33}$$

$$\limsup_{n \to \infty} \| T^n \vec{v}_0 \|^{1/n} \leqq \max \{ | \lambda | : \lambda \epsilon \sigma_2 \} < R_\sigma . \tag{34}$$

Therefore, the asymptotic behavior of $\vec{x}_n = \mathbf{T}^n \vec{x}_0$ as $n \to \infty$ is essentially that of $\mathbf{T}^n \vec{u}_0$ if $\vec{u}_0 \neq 0$. This reduces the problem to one in a *finite dimensional* subspace of B^m.

The condition above that $\vec{u}_0 \neq 0$ is not essential. If, instead, $\vec{x}_0$ has a non zero component in the spectral subspace M_λ^ν for some eigenvalue λ, then a modification of the foregoing argument with R_σ replaced by the maximum modulus of all such λ yields similar results. Further details are omitted.

The simplest special case is: $P(\lambda) \neq 0$ for $|\lambda| = R_\sigma$; there is just one eigenvalue λ_1 such that $|\lambda_1| = R_\sigma$; $\nu(\lambda_1) = 1$; and $\vec{u}_0 \neq 0$. Then $\mathbf{T}\vec{u}_0 = \lambda_1 \vec{u}_0$ and, hence,

$$\vec{x}_n = \mathbf{T}^n \vec{x}_0 = \lambda_1^n \vec{u}_0 + \mathbf{T}^n \vec{v}_0 , \tag{35}$$

where $\lambda_1^n \vec{u}_0$ is the asymptotically dominant term on the right. It follows from (15) that $T\vec{u}_0 = \lambda_1 \vec{u}_0$ if and only if

$$\vec{u}_0 = (u_0 , \lambda_1 u_0 , \dots , \lambda_1^{m-1} u_0) , \tag{36}$$

and

$$\left(\lambda_1^m I - \sum_{j=1}^{m} \lambda_1^{m-j} T_j) u_0 = 0 . \right) \tag{37}$$

Since T_j is of the form $T_j = a_j I + K_j$, it follows from (25) that (37) is equivalent to

$$[P(\lambda_1) I - \sum_{j=1}^{m} \lambda_1^{m-j} K_j] u_0 = 0 . \qquad (38)$$

Since $P(\lambda_1) \neq 0$ by hypothesis, this is a generalized Fredholm equation of the second kind. The number λ_1 is an eigenvalue of $\mathbf{T}$ if and only if (38) has a non-zero solution u_0, in which case (36) gives a corresponding eigenvector $\vec{u}_0$.

A special case of a composite recursion relation was studied by D. Greenspan and the author in [4]. Asymptotic results were obtained there which go beyond those given above.

We conclude this paper with a generalization of Theorem 1. Let $\mathfrak{A}$ be an algebra with unit I over the complex field. Let $\mathscr{I}$ be an ideal in $\mathfrak{A}$. Let $\mathfrak{A}_m$ denote the algebra of all $m \times m$ matrices $\mathbf{T} = [T_{ij}]$ with $T_{ij} \in \mathfrak{A}$. Then the set

$$\mathscr{I}_m = \{\mathbf{K} = [K_{ij}] : K_{ij} \in \mathscr{I}\} \qquad (39)$$

is an ideal in $\mathfrak{A}_m$.

Theorem 2. Let $\mathbf{T} = [a_{ij} I + K_{ij}]$, where $K_{ij} \in \mathscr{I}$. Let $P(\lambda)$ be the characteristic polynomial of the scalar matrix $[a_{ij}]$. Then $P(\mathbf{T}) \in \mathscr{I}_m$.

Since the proof is essentially the same as that for Theorem 1, it is omitted.

REFERENCES

[1] Taylor, A. E., *An Introduction to Functional Analysis*, John Wiley and Sons, New York, 1958.

[2] Doob, J. L., *Stochastic Processes*, John Wiley and Sons, New York, 1953.

[3] Coddington, E. A. and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill Book Co., New York, 1955.

[4] Anselone, P. M. and D. Greenspan, On a Class of Linear Difference-Integral Equations, Archive for Rat. Mech. and Anal. (to appear), and Math. Research Center Report 292, Madison, Wisc., 1962.

Mathematics Research Center
of the United States Army
University of Wisconsin
Madison, Wisconsin.