JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematics and Mathematical Physics
Vol. 64B, No. 1, January-March 1960

Space of k-Commutative Matrices[1]

Marvin Marcus[2] and N. A. Khan[3]

(August 14, 1959)

Let $[A, X]=AX-XA$ and $[A, X]_k=[A, [A, X]_{k-1}]$. Those matrices X which "k-commute" with a fixed matrix A are investigated. In particular, the dimension of the null space of the linear transformation $T(X)=[A, X]_k$ when A is nonderogatory is determined.

1. Introduction

Let A be a fixed N-square complex matrix and let $[A, X]=AX-XA$, $[A, X]_k=[A, [A, X]_{k-1}]$. It is easily checked that

$$[A,X]_k=\sum_{s=0}^{k}(-1)^s\binom{k}{s}A^{k-s}XA^s. \tag{1.1}$$

It is clear that the set of X such that $[A, X]_k=0$ is a linear subspace of the space M_N of all N-square complex matrices. This subspace is denoted by $\mathfrak{A}_k(A)$. In theorem 1 is determined the dimension of $\mathfrak{A}_k(A)$ in terms of the degrees of the elementary divisors of A when there is exactly one elementary divisor for each eigenvalue. Let E_q denote the set of matrices in M_N with precisely q distinct eigenvalues. In theorem 2 it is shown that in case $k\geq 2(N-q)+1$, then

$$\min_{A\epsilon E_q\cap M_N}\dim \mathfrak{A}_k(A)=R(Q+1)^2+(q-R)Q^2$$

where $N=Qq+R$, $0\leq R<q$. The maximum is also found.

2. Results

Let T denote the linear transformation on M_N defined by $T(X)=[A, X]$ and we note that $T^k(X)=[A, X]_k$. With respect to the basis E_{ij} in M_N, ordered lexicographically, we check that T has the matrix representation $I_N\otimes A-A'\otimes I_N$. The notation is the following: E_{ij} is the N-square matrix with 1 in position i, j, 0 elsewhere; $I_N\otimes A$ denotes the Kronecker product of the N-square identity matrix I_N with A. It is clear that one may assume A is in Jordan canonical form

$$A=\sum_{s=1}^{q}{}^{\bullet}\, J_s, \tag{2.1}$$

where $\Sigma^{\bullet}$ indicates direct sum and the J_s are the Jordan blocks corresponding to the distinct eigenvalues λ_s, $s=1, \ldots, q$, of A. If $T_k(X)=0$ and X is partitioned conformally with the partitioning (2.1) of A, Roth[4] shows that

$$X=\sum_{s=1}^{q}{}^{\bullet}\, X_s,$$

where X_s is the same size as J_s, $s=1, \ldots, q$. If, further, each J_s is decomposed into a direct sum of companion matrices of the elementary divisors corresponding to λ_s one may also effect a conformal partitioning of the corresponding X_s. It is also clear that one may take $\lambda_s=0$ in examining the structure of X_s since $[J_s, X_s]_k$ remains invariant upon translation of J_s. We are thus reduced to considering the following situation in determining the dimension of $\mathfrak{A}_k(A)$: Let

$$U=\sum_{i=1}^{l}{}^{\bullet}\, U_{r_i} \qquad r_1\geq \ldots \geq r_l,$$

where U_{r_i} is an r_i-square auxiliary unit matrix, an unbroken line of 1's along the first super-diagonal 0's elsewhere, and suppose $[U, Y]_k=0$. Partition Y conformally with U, $Y=(Y_{ij})$ $i,j=1, \ldots, l$ where Y_{ij} is $r_i\times r_j$, and by (1.1)

$$\sum_{s=0}^{k}(-1)^s\binom{k}{s}U_{r_i}^{k-s}Y_{ij}U_{r_j}^{s}=0 \qquad i, j=1, \ldots, l \tag{2.2}$$

is equivalent to $[U, Y]_k=0$. The problem then is to determine the number of arbitrary parameters in each Y_{ij}.

Equation (2.2) for a fixed i, j represents a linear transformation mapping Y_{ij} into 0, and with respect to a suitably chosen basis this transformation has the matrix representation T_{ij}^k where

$$T_{ij}=(I_{r_i}\otimes U_{r_j}-U'_{r_i}\otimes I_{r_j}). \tag{2.3}$$

To simplify the notation, put $r_i=n$, $r_j=m$ where it can be assumed without loss of generality that $n\geq m$. The similarity invariants of T_{ij} as computed by

[1] This work was supported in part by U.S. National Science Foundation Grant NSFG-5416.
[2] Present address: The University of British Columbia, Vancouver, Canada.
[3] Present address: Muslim University, Aligarh, India.

[4] W. E. Roth, On k-commutative matrices, Trans. Am. Math. Soc. **39**, 483 (1936).

Roth[5] are $f_1(x)= \dots =f_{mn-m}(x)=1$, $f_{mn-m+p}(x)=x^{\Delta+2p-1}$, $p=1, \dots, m$ and $\Delta=n-m$. Hence T_{ij} is similar to the direct sum of the companion matrices of these nontrivial similarity invariants,

$$T_{ij}\simeq\sum_{p=1}^{m}{}^{\bullet}\ C(x^{\Delta+2p-1}). \tag{2.4}$$

The sizes of these companion matrices arranged in decreasing order are $\Delta+2m-1$, $\Delta+2m-3, \dots,$ $\Delta+3$, $\Delta+1$.

Now, if $k\geq\Delta+2p-1$, then $(C(x^{\Delta+2p-1}))^k=0$. If $k<\Delta+2p-1$, then

$$\rho(\{C(x^{\Delta+2p-1})\}^k)=\Delta+2p-1-k,$$

where ρ denotes rank.

Let η denote nullity.

Lemma 1: (a) $\eta(T_{ij}^k)=km$ *if* $1\leq k<\Delta$,

(b) $\eta(T_{ij}^k)=km-\left(\frac{k-\Delta}{2}\right)^2+C$ *if* $\Delta\leq k<m+n-1$,

(c) $\eta(T_{ij}^k)=mn$ *if* $k\geq m+n-1$,

where C is 0 or 1/4 according as $(k-\Delta)$ is even or odd.

Proof:

(a) $1\leq k<\Delta$.

Then

$$\rho(T_{ij}^k)=\sum_{p=1}^{m}(\Delta+2p-1-k)=mn\text{-}mk,$$

and

$$\eta(T_{ij}^k)=mk.$$

(b) $\Delta\leq k<n+m-1$.

Assume $k-\Delta$ is odd and observe that the size of the $(m-(k-\Delta+1)/2)$th companion matrix in (2.4) is

$$\Delta+2m-\left(2\left(m-\frac{(k-\Delta+1)}{2}\right)-1\right)=k+2,$$

and the size of the next companion matrix is k. Hence,

$$\rho(T_{ij}^k)=(\Delta+2m-1-k)+(\Delta+2m-3-k)+\dots+(k+2-k).$$

The last term in this sum is the rank of the kth power of the $(m-(k-\Delta+1)/2)$th companion matrix in (2.4). Thus,

$$\rho(T_{ij}^k)=\left(\frac{2m-k+\Delta-1}{2}\right)\left(\frac{2m-k+\Delta+1}{2}\right)$$
$$=\left(m-\frac{(k-\Delta)}{2}\right)^2-\frac{1}{4}.$$

In case $k-\Delta$ is even, it is observed that the size of the $(m-(k-\Delta)/2)$th companion matrix in (2.4) is $k+1$. Also the size of the next companion matrix is $k-1$. Hence,

$$\rho(T_{ij}^k)=(\Delta+2m-1-k)+(\Delta+2m-3-k)+\dots+(k+1-k)$$
$$=\left(m-\frac{(k-\Delta)}{2}\right)^2.$$

Hence, in either case

$$\eta(T_{ij}^k)=mn-\rho(T_{ij}^k)$$
$$=mn-\left(m-\frac{(k-\Delta)}{2}\right)^2+C$$
$$=mk-\left(\frac{k-\Delta}{2}\right)^2+C$$

where C is 0 or 1/4 depending on whether $k-\Delta$ is even or odd.

(c) $k\geq m+n-1$.

Then

$$\{C(x^{\Delta+2m-1})\}^k=0$$

and

$$\eta(T_{ij}^k)=mn.$$

Theorem 1. *Assume A is N-square with distinct eigenvalues $\lambda_1, \dots, \lambda_q$ and let $(x-\lambda_j)^{\epsilon_j}$ be the elementary divisors of A, $j=1, \dots, q$, $\epsilon_1\geq\epsilon_2\geq\dots\geq\epsilon_q$. Partition the integers $1, \dots, q$ so that*

$$\epsilon_1=\dots=\epsilon_{q_1}>\epsilon_{q_1+1}=\dots=\epsilon_{q_2}>\dots>\epsilon_{q_{t-1}+1}=\dots=\epsilon_{q_t}=\epsilon_q.$$

Then

(i) $\dim \mathfrak{A}_k(A)=kN-q\left(\frac{k^2}{4}-C\right)$ *if* $k<2\epsilon_q-1$,

(ii) $\dim \mathfrak{A}_k(A)=k\sum_{j=1}^{q_{t-1}}\epsilon_j+\sum_{j=q_{t-1}+1}^{q}\epsilon_j^2-q_{t-1}\left(\frac{k^2}{4}-C\right)$ *if* $2\epsilon_{q_t}-1\leq k<2\epsilon_{q_{t-1}}-1$,

(iii) $\dim \mathfrak{A}_k(A)=\sum_{j=1}^{q}\epsilon_j^2$ *if* $k\geq2\epsilon_1-1$, *where C is 0 or ¼ according as k is even or odd.*

[5] W. E. Roth, On direct product matrices, Bull. Am. Math. Soc. **40**, 461 (1934).

PROOF: Since there is only one elementary divisor of A for each eigenvalue of A it may be assumed, as in (2.1), that

$$A=\sum_{s=1}^{q}{}^{\bullet}\ J_s,$$

where $J_s=\lambda_s I_{\epsilon_s}+U_{\epsilon_s}$, U_{ϵ_s} the ϵ_s-square auxiliary unit matrix, $s=1, \ldots, q$. Then, if X is contained in $\mathfrak{A}_k(A)$,

$$X=\sum_{s=1}^{q}{}^{\bullet}\ X_s,$$

where X_s is ϵ_s-square, $s=1, \ldots, q$. By lemma 1 we know that the number of arbitrary parameters in X_s is (by putting $m=n=\epsilon_s$)

$$(b)\qquad n_s=\epsilon_s k-\frac{k^2}{4}+C \qquad \text{if} \quad k<2\epsilon_s-1,$$

$$(c)\qquad \epsilon_s^2 \qquad \text{if} \quad k\geq 2\epsilon_s-1.$$

Consider (*i*) first: $k<2\epsilon_q-1$. Then $k<2\epsilon_{q_j}-1$ $j=1, \ldots, l$, and hence X_s has n_s arbitrary parameters in it. There are $q_\sigma-q_{\sigma-1}$ values of s such that $n_s=n_{q_\sigma}$ ($q_0=0$ for convenience). Hence,

$$\begin{aligned}\dim \mathfrak{A}_k(A)&=\sum_{s=1}^{q} n_s=\sum_{\sigma=1}^{l} (q_\sigma-q_{\sigma-1})n_{q_\sigma}\\ &=\sum_{\sigma-1}^{l} (q_\sigma-q_{\sigma-1})\left(k\epsilon_{q_\sigma}-\frac{k^2}{4}+C\right)\\ &=k\sum_{\sigma-1}^{l} (q_\sigma-q_{\sigma-1})\epsilon_{q_\sigma}-q\left(\frac{k^2}{4}-C\right)\\ &=k\sum_{j=1}^{q} \epsilon_j-q\left(\frac{k^2}{4}-C\right)\\ &=kN-q\left(\frac{k^2}{4}-C\right).\end{aligned}$$

Next assume that (ii) $2\epsilon_{q_t}-1\leq k<2\epsilon_{q_{t-1}}-1$.

In this case $n_{q_\sigma}=\epsilon_{q_\sigma}^2$, $\sigma=t, \ldots, l$ and $n_{q_\sigma}=k\epsilon_{q_\sigma}-\frac{k^2}{4}+C$, $\sigma=1, \ldots, t-1$. Hence,

$$\begin{aligned}\dim \mathfrak{A}_k(A)&=\sum_{\sigma-1}^{t-1} (q_\sigma-q_{\sigma-1})n_{q_\sigma}+\sum_{\sigma=t}^{l} (q_\sigma-q_{\sigma-1})\epsilon_{q_\sigma}^2\\ &=k\sum_{s-1}^{q_{t-1}} \epsilon_s+\sum_{s=q_{t-1}+1}^{q} \epsilon_s^2-q_{t-1}\left(\frac{k^2}{4}-C\right).\end{aligned}$$

If (*iii*) $k\geq 2\epsilon_1-1$, then $k\geq 2\epsilon_j-1$ for $j=1, \ldots, q$ and

$$\dim \mathfrak{A}_k(A)=\sum_{s=1}^{q}\epsilon_s^2.$$

In case there is more than one elementary divisor corresponding to a particular eigenvalue there does not seem to be any simple formula for dim $\mathfrak{A}_k(A)$ in terms of the degrees of the elementary divisors of A. However, by repeated use of lemma 1, it is possible to compute dim $\mathfrak{A}_k(A)$ for any particular A. For example, if

$$A=\begin{pmatrix}2&1&0&0\\0&2&1&0\\0&0&2&1\\0&0&0&2\end{pmatrix}\dot{+}\begin{pmatrix}2&1\\0&2\end{pmatrix}\dot{+}(2),$$

then $\dim \mathfrak{A}_k(A)=37$ for $k=3$.

Next are determined the largest and smallest values that dim $\mathfrak{A}_k(A)$ may take on as A varies over E_q, the set of matrices with precisely q distinct eigenvalues, under the condition that $k\geq 2(N-q)+1$.

LEMMA 2. *If* $\epsilon_1\geq \ldots \geq\epsilon_q$ *are positive integers satisfying*

$$\sum_{j=1}^{q}\epsilon_j=N=qQ+R, \qquad 0\leq R<q,$$

then

$$R(Q+1)^2+(q-R)Q^2\leq \sum_{j=1}^{q}\epsilon_j^2\leq (N-q+1)^2+(q-1).$$

The lower bound is achieved for

$$\epsilon_1= \ldots =\epsilon_R=Q+1, \qquad \epsilon_{R+1}= \ldots =\epsilon_q=Q,$$

and the upper bound is achieved for

$$\epsilon_1=N-q+1 \quad \textit{and} \quad \epsilon_2= \ldots =\epsilon_q=1.$$

PROOF. The lower inequality is proved by induction on R. In case $R=0$, then $N/q=Q$, and if $\epsilon_1, \ldots, \epsilon_q$ are regarded as continuous variables, then $\sum_{j=1}^{q}\epsilon_j^2$ has a minimum for $\epsilon_j=Q$, $j=1, \ldots, q$. Now suppose the result is true for all remainders obtained by dividing N by q that are less than R. We first claim that there exists an integer $i<q$ such that $\epsilon_i>\epsilon_{i+1}$ and $\epsilon_i\geq Q+1$. Clearly the set of integers j such that $\epsilon_j\geq Q+1$ is nonempty otherwise (since $R>0$),

$$\sum_{j=1}^{q}\epsilon_j\leq qQ<N.$$

Let i be the largest integer j such that $\epsilon_j\geq Q+1$; then if i were q, $\epsilon_1\geq \ldots \geq\epsilon_q\geq Q+1$ and

$$\sum_{j=1}^{q}\epsilon_j\geq q(Q+1)>N.$$

Hence $i<q$ and from the definition of i, $\epsilon_i>\epsilon_{i+1}$. Let $\mu_j=\epsilon_j$, $j\neq i$ and $\mu_i=\epsilon_i-1$. Then $\mu_1\geq \ldots \geq\mu_q$, $\mu_1+\ldots+\mu_q=N-1=Qq+R-1$, and by induction

$$\sum_{j=1}^{q}\mu_j^2\geq (R-1)(Q+1)^2+(q-(R-1))Q^2=R(Q+1)^2 +(q-R)Q^2+(Q^2-(Q+1)^2).$$

Now

$$\sum_{j=1}^{q}\mu_j^2=\sum_{j=1}^{q}\epsilon_j^2-2\epsilon_i+1,$$

and thus,

$$\sum_{j=1}^{q}\epsilon_j^2\geq R(Q+1)^2+(q-R)Q^2+2(\epsilon_i-Q-1).$$

Since $\epsilon_i\geq Q+1$, the proof is complete. The upper bound is easily obtained.

THEOREM 2. *If* $k\geq 2(N\text{-}q)+1$, *then*

$$\min_{A\epsilon E_q\cap M_N}\dim \mathfrak{A}_k(A)=R(Q+1)^2+(q-R)Q^2$$

and

$$\max_{A\epsilon E_q\cap M_N}\dim \mathfrak{A}_k(A)=(N-q+1)^2+q-1,$$

where

$$N=qQ+R,\qquad 0\leq R<q.$$

PROOF. Let $A\epsilon E_q$ and suppose $\epsilon_1\geq \ldots \geq\epsilon_q$ are such integers that ϵ_j is the sum of the degrees of all elementary divisors of A corresponding to λ_j, $j=1, \ldots, q$.

Then

$$\sum_{=1}^{q}\epsilon_j=N,$$

and hence, $\epsilon_1\leq N-q+1$ and $2\epsilon_1-1\leq k$. Thus k is at least $2\mu-1$ where μ is the degree of any elementary divisor of A. From lemma 1 one may check in this case that $\min_{A\epsilon E_q\cap M_N}\dim \mathfrak{A}_k(A)$ may be evaluated by confining A to those matrices having precisely one elementary divisor for each eigenvalue. Hence $(x-\lambda_j)\epsilon_j$ may be taken as the elementary divisors of A, $j=1, \ldots, q$. By theorem 1 if $A\epsilon E_q$,

$$\dim \mathfrak{A}_k(A)=\sum_{j=1}^{q}\epsilon_j^2,$$

and the results follow from lemma 2.

WASHINGTON, D.C. (Paper 64B1–21)