A NOTE ON MATRIX COMMUTATORS

D. W. Robinson

The problem of representing a given matrix as a matrix commutator has received attention from several authors. (See for example [1], [5], [6], or [7].) Motivated by a recent paper of I. N. Herstein [2], this note provides necessary and sufficient conditions for representing a given nonsingular matrix as a multiplicative commutator $ABA^{-1}B^{-1}$ such that the additional condition $A(AB - BA) = (AB - BA)A$ is satisfied.

LEMMA. *Let* D *be a nonsingular* n-*by*-n *matrix over a field* F *of characteristic zero or prime* $p > n$. *Then there exist nonsingular matrices* A *and* B *over* F *such that* $D = ABA^{-1}B^{-1}$ *and* $A(AB - BA) = (AB - BA)A$ *if and only if* $D - I$ *is nilpotent.*

Necessity. The necessity of this condition is a restatement of a theorem proved by C. R. Putnam and A. Wintner [3] for fields of characteristic zero and by I. N. Herstein [2] for fields of prime characteristic $p > n$.

Sufficiency. Let $D - I$ be nilpotent. Since $D - I$ is similar to a direct sum of matrices, each nilpotent of index equal to its order, it is clearly sufficient to prove the result for the n-by-n matrix $D - I$ that is nilpotent of index n. Furthermore, without loss of generality, let $D = (d_{ij})$, where $d_{ij} = \delta_{ij} + \delta_{i,j-1}$, be in classical canonical form.

Let $A = (a_{ij})$, where $a_{ij} = \binom{j}{i}$ for $i \leq j$ and $a_{ij} = 0$ otherwise. By matrix multiplication and by the addition properties of binomial coefficients, it is easily shown that $A(2I - D) = (a_{i-1,j-1})$ and that $DA(2I - D) = A$. Hence $D^{-1}A = A(2I - D)$ and $A(D - I) = (I - D^{-1})A$. By the restriction on the characteristic of the field, it is evident that none of the elements $a_{k-1,k} = k$ of the first superdiagonal of either A or $D^{-1}A$ are zero. Hence, both $A - I$ and $D^{-1}A - I$ are nilpotent of index n, and this implies that A and $D^{-1}A$ are similar. Thus, there is a nonsingular matrix B such that $D^{-1}A = BAB^{-1}$, and since A is nonsingular, $D = ABA^{-1}B^{-1}$. Finally,

$$A(AB - BA) = A(ABA^{-1}B^{-1} - I)BA = A(D - I)BA = (I - D^{-1})ABA$$

$$= (I - BAB^{-1}A^{-1})ABA = (AB - BA)A.$$

The preceding proof suggests the following theorem.

THEOREM. *Let* D *be a nonsingular* n-*by*-n *matrix over a field* F *other than the field of two elements. Then there exist nonsingular matrices* A *and* B *over* F *such that* $D = ABA^{-1}B^{-1}$ *and* $A(AB - BA) = (AB - BA)A$ *if and only if* $D - I$ *is similar to* $I - D^{-1}$.

Necessity. Let $D = ABA^{-1}B^{-1}$ and $A(AB - BA) = (AB - BA)A$. Then

$$A(D - I) = A(ABA^{-1}B^{-1} - I) = A(AB - BA)A^{-1}B^{-1} = (AB - BA)AA^{-1}B^{-1}$$

$$= (AB - BA)B^{-1}A^{-1}A = (I - BAB^{-1}A^{-1})A = (I - D^{-1})A.$$

Received September 21, 1959.

Since A is nonsingular, it follows that $D - I$ is similar to $I - D^{-1}$.

Sufficiency. $D - I$ is similar to the direct sum of a nonsingular matrix, say $D_1 - I_1$, and a nilpotent matrix, say $D_2 - I_2$. Since the same similarity transformation gives $I - D^{-1} = (D - I)D^{-1}$ as the direct sum of the nonsingular matrix $I_1 - D_1^{-1}$ and the nilpotent matrix $I_2 - D_2^{-1}$, under the hypothesis that $D - I$ is similar to $I - D^{-1}$, the corresponding matrices of the direct sums are also similar. Hence, it is sufficient to prove the result for the following two cases.

First, let $D - I$ be nonsingular, and define $B = (D - I)^{-1}$. Since $D - I$ is similar to $I - D^{-1}$, there is a nonsingular A such that $A(D - I) = (I - D^{-1})A$. Hence,

$$AB^{-1} = A(D - I) = (I - D^{-1})A = (D - I)D^{-1}A = B^{-1}D^{-1}A.$$

That is, $D = ABA^{-1}B^{-1}$. Also, by an argument analogous to that given in the proof of the lemma, $A(AB - BA) = (AB - BA)A$.

Second, let $D - I$ be nilpotent. By the lemma above, it is only necessary to prove the result for fields of prime characteristic $p \leq n$. Also, consider again $D - I$ to be nilpotent of index n and to be in classical canonical form.

Consider first $p \neq 2$, and choose r and s such that $n - 1 = pr + 2s$, where $0 \leq s < p$ and $r \geq 0$. For convenience, let $U = (u_{ij})$ be the $(r + 2)p$-by-$(r + 2)p$ matrix with elements $u_{ij} \equiv \binom{j-1}{i-1}$ (modulo p) for $i \leq j$ and zero otherwise. By induction on the positive integer t, it is easily verified that $U^t = (t^{j-i}u_{ij})$. Hence, in particular, $U^p = I$ and $(U - I)^p = 0$. That is, $U - I$ is nilpotent of index at most p. On the other hand, since the first superdiagonal of U has nonzero elements in $(r + 2)(p - 1)$ positions, the nullity of $U - I$ is at most $r + 2$. Hence, since the order of U is $(r + 2)p$, it follows that $U - I$ is similar to a direct sum of $r + 2$ nilpotent matrices each of index p.

Now, let A be the principal submatrix of U obtained by deleting the first $p - s$ and the last $p - (s + 1)$ rows and columns. Thus, $A = (u_{i+p-s,j+p-s})$ is an n-by-n matrix. As in the proof of the lemma, $DA(2I - D) = A$ and $D^{-1}A = (u_{i+p-s-1,j+p-s-1})$. That is, $D^{-1}A$ is the principal submatrix of U obtained by deleting the first $p - (s + 1)$ and the last $p - s$ rows and columns. Clearly, both $A - I$ and $D^{-1}A - I$ are similar to a direct sum of $r + 2$ ($r + 1$ if $s = 0$) nilpotent matrices, r of index p, one of index $s + 1$, and one of index s. Hence, A is similar to $D^{-1}A$, and the conclusion follows as before.

Finally, let $p = 2$. If n is odd, then the preceding argument is valid with the choices $s = 0$ and $r = (n - 1)/2$. If $n = (r + 2)2$ is even with $r \geq 0$, then let $A = U + M$, where U is given above and $M = (m_{ij})$ is the n-by-n matrix with all elements zero except $m_{1n} = 1$. Since $DMD = M$, it follows that $DA(2I - D) = DAD = A$. Furthermore, since $U - I$, $D^{-1}U - I$, M, and $D^{-1}M$ are all nilpotent of index 2, the same is true of $A - I$ and $D^{-1}A - I$. Also, it is evident that the nullity of both $A - I$ and $D^{-1}A - I$ is $r + 2$. Hence, both $A - I$ and $D^{-1}A - I$ are similar to a direct sum of $r + 2$ matrices, each nilpotent of index 2. Again the conclusion follows.

The preceding argument is not valid for $n = 2$, the only remaining case to be considered. Indeed, no construction is possible over the field of two elements, for it is easily demonstrated that

$$D = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

cannot be represented as a multiplicative commutator over this field. (See [7].) However, if F is not the field of two elements, then, with the choices $\alpha \neq 0, 1$, and

$$A = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} \alpha + 1 & 0 \\ 0 & \alpha \end{pmatrix},$$

the desired representation is obtained. This result completes the proof of the theorem.

Two supplementary remarks are in order. First, it is evident that any matrix D over the field of two elements with an eigenvalue 1 of multiplicity 2 and index 2 cannot be represented as a commutator $ABA^{-1}B^{-1}$ with $A(AB - BA) = (AB - BA)A$. It can be shown, moreover, that a matrix of this type is the only exception, provided of course that $D - I$ is similar to $I - D^{-1}$.

Second, it is clear that if $D - I$ is nilpotent, then $D - I$ is similar to $I - D^{-1}$. By the results above, the converse is true for fields of characteristic zero or prime $p > n$. However, for fields of prime characteristic $p \leq n$, it can be shown that the converse is not valid. (See [4].) Hence, in particular, the restriction on the characteristic of the field in the lemma above cannot be relaxed.

REFERENCES

1. A. A. Albert and B. Muckenhoupt, *On matrices of trace zero,* Michigan Math. J. 4 (1957), 1-3.

2. I. N. Herstein, *On a theorem of Putnam and Wintner,* Proc. Amer. Math. Soc. 9 (1958), 363-364.

3. C. R. Putnam and A. Wintner, *On the spectra of group commutators,* Proc. Amer. Math. Soc. 9 (1958), 360-362.

4. D. W. Robinson, *On a theorem arising from a study of matrix commutators,* Abstract 556-12, Notices Amer. Math. Soc. 6 (1959), 166.

5. K. Shoda, *Einige Sätze über Matrizen,* Jap. J. Math. 13 (1936), 361-365.

6. ———, *Über den Kommutator der Matrizen,* J. Math. Soc. Japan 3 (1951), 78-81.

7. R. C. Thompson, *On matrix commutators,* Abstract 556-7, Notices Amer. Math. Soc. 6 (1959), 164.

Brigham Young University