A CONCEPTUAL APPROACH TOWARDS UNDERSTANDING MATRIX COMMUTATORS

SOUMYASHANT NAYAK

Abstract. It is well known that a square matrix over a field is a commutator if and only if it has trace zero. We give a conceptual proof of this result for square matrices over algebraically closed fields.

Keywords: Commutators, Trace, Similarity
MSC2010 subject classification: 15A18, 15A21

1. Introduction

Let $\mathbb{K}$ be a field. For a natural number n, let $M_n(\mathbb{K})$ denote the $\mathbb{K}$-algebra of $n \times n$ matrices with entries from $\mathbb{K}$. The *commutator* of two matrices $P, Q \in M_n(\mathbb{K})$ is defined as $[P, Q] := PQ - QP$. A *trace* τ on $M_n(\mathbb{K})$ is a $\mathbb{K}$-valued linear map on $M_n(\mathbb{K})$ such that $\tau(PQ) = \tau(QP)$ for all $P, Q \in M_n(\mathbb{K})$. In other words, a trace on $M_n(\mathbb{K})$ is a $\mathbb{K}$-valued linear map on $M_n(\mathbb{K})$ that vanishes on the set of commutators in $M_n(\mathbb{K})$. There is a $\mathbb{K}$-valued trace (unique up to scalar multiplication) on $M_n(\mathbb{K})$, which is given by the sum of the diagonal entries for a given matrix.

In [2], Shoda showed that an element of $M_n(\mathbb{K})$ with trace zero is necessarily a commutator when $\mathbb{K}$ has characteristic zero. The result for arbitrary fields was shown by Albert and Muckenhoupt in [1]. The present author is of the opinion that the proofs available in the literature lean more towards the technical side and do not throw as much light on why we should expect such a result to hold. In this note, we provide a conceptual proof of the fact that trace zero matrices over an algebraically closed field are commutators.

2. Trace zero matrices and commutators

In this section, $\mathbb{K}$ denotes an algebraically closed field. For $X, Y \in M_n(\mathbb{K})$, we say that X is *similar* to Y if there is an invertible matrix $S \in M_n(\mathbb{K})$ such that $X = SYS^{-1}$. It is easy to see that the relation, *similarity*, is an equivalence relation on $M_n(\mathbb{K})$. The equivalence class of an element $X \in M_n(\mathbb{K})$ is said to be the *similarity orbit* of X.

Let the matrix $X \in M_n(\mathbb{K})$ be a commutator, so that there are matrices $P, Q \in M_n(\mathbb{K})$ such that $X = [P, Q]$. Note that $[P, Q] = [P - \alpha I, Q]$ for every $\alpha \in \mathbb{K}$. Since $\mathbb{K}$ has infinitely many elements (by virtue of being algebraically closed), there exists an $\alpha \in \mathbb{K}$ such that $P - \alpha I$ is invertible and hence without loss of generality, we may assume that P is invertible. Let us rewrite X as $P(QP)P^{-1} - QP$ and define $T := QP$. Thus $PTP^{-1} = T + X$ or in other words, T and $T + X$ are similar. Conversely, it is easy to see by retracing our steps that if there is a matrix $T \in M_n(\mathbb{K})$ such that T and $T + X$ are similar, then X is a commutator. Also note that if $X = [P, Q]$, then $SXS^{-1} = [SPS^{-1}, SQS^{-1}]$ for every invertible matrix $S \in GL_n(\mathbb{K})$. We summarize the preceding discussion in the following lemma.

Lemma 2.1. *For a matrix $X \in M_n(\mathbb{K})$, the following are equivalent:*

(i) *X is a commutator;*
(ii) *There is a matrix in the similarity orbit of X which is a commutator;*
(iii) *There is a matrix $T \in M_n(\mathbb{K})$ such that T and $T + X$ are similar.*

In view of Lemma 2.1 and recalling that $\mathbb{K}$ is assumed to be algebraically closed, we will take the liberty of viewing X in its Jordan canonical form. In this note, our main goal is to show the following result, which combined with Lemma 2.1, shows that trace zero matrices over an algebraically closed field are commutators.

Theorem 2.2. *Let X be a matrix in $M_n(\mathbb{K})$. Then there is a matrix $T \in M_n(\mathbb{K})$ such that T and $T+X$ are similar if and only if X has trace zero.*

If T and $T+X$ are similar, they must have the same trace and hence it easily follows that X must have trace zero. A key element of our proof of the other direction is the following basic linear algebra result.

Proposition 2.3. *Let $\mathbb{K}$ be an algebraically closed field. If every eigenvalue of a matrix $A \in M_n(\mathbb{K})$ has multiplicity 1, then A is diagonalizable. Consequently, if two matrices $A, B \in M_n(\mathbb{K})$ have the same set of eigenvalues with each eigenvalue having multiplicity 1, then A and B are similar.*

Note that nilpotent matrices in $M_n(\mathbb{K})$ have trace zero. We discuss Theorem 2.2 for the special case of nilpotent matrices before proceeding to the general case. Let $X \in M_n(\mathbb{K})$ be a nilpotent matrix in Jordan canonical form so that it is an upper triangular matrix with zeroes on the principal diagonal. Since the field $\mathbb{K}$ has infinitely many elements, we may choose distinct elements $\lambda_1, \lambda_2, \ldots, \lambda_n$ from $\mathbb{K}$. Let $D := \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Note that D and $D+X$ have the same set of eigenvalues. By Proposition 2.3, D and $D+X$ are similar. By Lemma 2.1, X is a commutator. We conclude that every nilpotent matrix in $M_n(\mathbb{K})$ is a commutator.

The proof of the main theorem (Theorem 2.2) involves use of the following combinatorial lemma. For the sake of brevity, we use the notation $[n] := \{1, 2, \ldots, n\}$ in the lemma. We denote the group of permutations of $[n]$ by Σ_n.

Lemma 2.4. *Let G be an abelian group with infinitely many elements and n be a positive integer. Let $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ be an n-tuple of elements from G satisfying*

$$\sum_{i=1}^{n} \lambda_i = 0.$$

Then there is an n-tuple $(\mu_1, \mu_2, \ldots, \mu_n)$ of elements from G and a permutation $\sigma \in \Sigma_n$ such that $\mu_i \neq \mu_j$ for $i \neq j$, and $\lambda_i = \mu_i - \mu_{\sigma(i)}$ for $1 \leq i \leq n$.

Proof. We proceed inductively. For $n = 1$, the lemma is obvious. Consider a positive integer $m \geq 2$, and assume that the assertion is true for $n = 1, 2, \ldots, m-1$.

Let $(\lambda_1, \lambda_2, \ldots, \lambda_m)$ be an m-tuple of elements from G satisfying $\sum_{i=1}^{m} \lambda_i = 0$. We say that a subset S of $[m]$ has property $\mathfrak{P}$ if $\sum_{i \in S} \lambda_i = 0$ and $\sum_{i \in S'} \lambda_i \neq 0$ for any non-empty proper subset, S', of S. Let R be a non-empty subset of $[m]$ with smallest cardinality such that $\sum_{i \in R} \lambda_i = 0$. Clearly R has property $\mathfrak{P}$.

<u>Case I</u>: $R = [m]$.

For $1 \leq k \leq m$, we define μ_k to be the cumulative sum $\sum_{i=1}^{k} \lambda_i$. Let σ be the n-cycle defined by

$$\sigma(i) = \begin{cases} i-1 & \text{if } 2 \leq i \leq m \\ m & \text{if } i = 1. \end{cases}$$

From the hypothesis of the lemma, $\mu_m = \sum_{i=1}^{m} \lambda_i = 0$. Furthermore, for $1 \leq k < \ell \leq m$, we have $\mu_\ell - \mu_k = \sum_{i=k+1}^{\ell} \lambda_k \neq 0$ (that is, $\mu_k \neq \mu_\ell$). For $k = 1$, we have $\mu_k - \mu_{\sigma(k)} = \mu_1 - \mu_{\sigma(1)} = \mu_1 - \mu_m = \lambda_1$. For $2 \leq k \leq m$, we observe that $\mu_k - \mu_{\sigma(k)} = \mu_k - \mu_{k-1} = \sum_{i=1}^{k} \lambda_i - \sum_{i=1}^{k-1} \lambda_i = \lambda_k$. Thus the m-tuple $(\mu_1, \mu_2, \ldots, \mu_m)$ satisfies the conditions described in the lemma.

Case II: R is a proper non-empty subset of $[m]$.

By suitable permutation of the entries of $(\lambda_1, \lambda_2, \ldots, \lambda_m)$, we may assume that $R = [\ell]$ for some $1 \leq \ell \leq m-1$. Note that $(\lambda_1, \ldots, \lambda_\ell)$ is an ℓ-tuple satisfying the hypothesis of the lemma. Using the induction hypothesis for $n = \ell$, we obtain an ℓ-tuple $(\mu_1, \mu_2, \ldots, \mu_\ell)$ and a permutation of $[\ell]$, σ_1, such that $\lambda_i = \mu_i - \mu_{\sigma_1(i)}$ for $1 \leq i \leq \ell$ and $\mu_i \neq \mu_j$ for distinct $i, j \in [\ell]$.

Note that
$$\sum_{i=\ell+1}^{m} \lambda_i = \sum_{i=1}^{m} \lambda_i - \sum_{i=1}^{\ell} \lambda_i = 0 - 0 = 0.$$
Thus $(\lambda_{\ell+1}, \ldots, \lambda_m)$ is an $(m-\ell)$-tuple satisfying the hypothesis of the lemma. Using the induction hypothesis for $n = m-\ell$, we obtain an $(m-\ell)$-tuple $(\mu_{\ell+1}, \ldots, \mu_m)$ and a permutation of $[m]\backslash[\ell]$, σ_2, such that $\lambda_i = \mu_i - \mu_{\sigma_2(i)}$ for $\ell+1 \leq i \leq m$ and $\mu_i \neq \mu_j$ for distinct $i, j \in [m]\backslash[\ell]$.

Let σ be the permutation of $[m]$ which acts as σ_1 on $[\ell]$ and as σ_2 on $[m]\backslash[\ell]$. Clearly $\lambda_i = \mu_i - \mu_{\sigma(i)}$ for $1 \leq i \leq m$. The set $\Lambda := \{\mu_i - \mu_j | i \in [\ell], j \in [m]\backslash[\ell]\}$ is a finite subset of G. Since G is infinite, we may choose $\alpha \in G\backslash\Lambda$. We define an m-tuple $(\nu_1, \nu_2, \ldots, \nu_m)$ as follows:
$$\nu_i := \begin{cases} \mu_i - \alpha & \text{if } i \in [\ell] \\ \mu_i & \text{if } \ell+1 \leq i \leq m \end{cases}$$
We observe that $\lambda_i = \nu_i - \nu_{\sigma(i)}$ for $1 \leq i \leq m$, and for distinct $i, j \in [m]$, we have $\nu_i \neq \nu_j$. Thus the m-tuple $(\nu_1, \nu_2, \ldots, \nu_m)$ satisfies the conditions described in the lemma. □

Proof of Theorem 2.2. As mentioned before, it is straightforward to see that X has trace zero if there is a matrix T such that T and $T + X$ are similar.

For the converse, without loss of generality, we may assume that X is in Jordan canonical form. Let $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ denote the n-tuple of eigenvalues of X such that
$$X = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) + N$$
where N is an upper triangular nilpotent matrix. Since $\operatorname{tr}(X) = \sum_{i=1}^{n} \lambda_i = 0$, from Lemma 2.4, we have an n-tuple $(\mu_1, \mu_2, \ldots, \mu_n)$ of elements of $\mathbb{K}$ such that $\mu_i \neq \mu_j$ for distinct $i, j \in [n]$ and a permutation $\sigma \in \Sigma_n$ such that $\lambda_i = \mu_i - \mu_{\sigma(i)}$. Let
$$D_\sigma := \operatorname{diag}(\mu_{\sigma(1)}, \mu_{\sigma(2)}, \ldots, \mu_{\sigma(n)}), \text{ and } D := \operatorname{diag}(\mu_1, \mu_2, \ldots, \mu_n).$$
By Proposition 2.3, D_σ and $D + N$ are similar as they have the same set of eigenvalues and each of their eigenvalues has multiplicity 1. Since $D + N = D_\sigma + X$, we conclude that D_σ and $D_\sigma + X$ are similar. Thus $T := D_\sigma$ satisfies the assertion in the theorem. □

From Lemma 2.1 and Theorem 2.2, we conclude that a square matrix over an algebraically closed field is a commutator if and only if it has trace zero.

References

[1] A. A. Albert and Benjamin Muckenhoupt. On matrices of trace zeros. *Michigan Math. J.*, 4:1–3, 1957.

[2] Kenjiro Shoda. Einige Sätze über Matrizen. *Jpn. J. Math.*, 13(3):361–365, 1937.

Indian Statistical Institute, Statistics and Mathematics Unit, 8th Mile, Mysuru Road, RVCE Post, Bengaluru, 560059, Karnataka, India

Email address: `soumyashant@isibang.ac.in, soumyashant@gmail.com`

URL: `https://nsoum.github.io/`