Matrices That Commute with Their Conjugate and Transpose

Geoffrey Goodson
Towson University, Towson, USA
Email: ggoodson@towson.edu

Received June 25, 2013; revised July 23, 2013; accepted August 6, 2013

Copyright © 2013 Geoffrey Goodson. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

It is known that if \( A \in M_n \) is normal \( (AA^* = A^*A) \), then \( A\bar{A} = \bar{A}A \) if and only if \( AA^T = A^TA \). This leads to the question: do both \( AA^* = A^*A \) and \( AA^T = A^TA \) imply that \( A \) is normal? We give an example to show that this is false when \( n = 4 \), but we show that it is true when \( n = 2 \) and \( n = 3 \).

Keywords: Normal Matrix; Matrix Commuting with Its Conjugate and Transpose

Introduction and Results

Let \( A \) be an \( n \times n \) normal matrix, i.e., \( A \) is a complex square matrix \( (A \in M_n) \), with the property that \( AA^* = A^*A \), where \( A^* = \bar{A}^T \) is the conjugate-transpose of \( A \). The Fuglede-Putnam Theorem tells us that if \( \bar{B} = \bar{A} \bar{B} \) for some \( \bar{B} \in M_n \), then \( \bar{A} \bar{B} = \bar{B} \bar{A} \). Suppose that \( A\bar{A} = \bar{A}A \), where \( A \) is the conjugate of the matrix \( A \) (so we take the complex conjugate of every entry of \( A \)). Then taking the transpose gives

\[
\bar{A}^T A^* = A^T \bar{A}^* \Rightarrow A^T A^* = A^T A^* \Rightarrow AA^T = A^TA,
\]

from the the Fuglede-Putnam Theorem. In a similar way, we see that if \( AA^* = A^*A \), then \( A\bar{A} = \bar{A}A \), so these two statements are equivalent when \( A \) is normal. The question arose in [2], whether the conditions

\[
\bar{A}A = A\bar{A} \quad \text{and} \quad A^T A^* = AA^T
\]

imply the third condition \( AA^* = A^*A \), so that \( A \) is normal.

This is false when \( n = 4 \). In fact, any matrix of the form

\[
A = \begin{bmatrix}
I_{ab} & I_{ac} \\
0 & I_{bd}
\end{bmatrix}, \quad \text{where} \quad I_{ab} = \begin{bmatrix}
a & b \\
b & -a
\end{bmatrix},
\]

\( a, b, c, d \in \mathbb{C} \), \( c^2 + d^2 = 0 \), \( c \) and \( d \) not both zero, has the property that both \( \bar{A}A = A\bar{A} \) and \( A^T A = AA^T \), but \( A \) is not normal. In this paper, we prove that if \( A \in M_n \) where \( n = 2 \) or \( n = 3 \), then these conditions do imply that \( A \) is normal. This result was first proposed as a problem by the current author in the International Linear Algebra Society journal IMAGE (fall 2011). My solution for \( n = 2 \) appeared in the spring 2012 issue, but no solution for the case \( n = 3 \) has ever been given. In this paper, we give the solution for the case \( n = 3 \), and for completeness, we also give the solution for \( n = 2 \). Specifically we prove:

**Theorem 1** If \( A \in M_n \), \( n = 2 \) or \( n = 3 \), then \( A\bar{A} = \bar{A}A \) and \( AA^T = A^TA \) imply that \( A \) is normal.

**Proof.** We need the following preliminary result, which is a direct consequence of Theorem 2.3.6 in [3] (using the fact that for \( A \in M_n \), \( A = B + iC \) where \( B \) and \( C \) are real then \( A\bar{A} = \bar{A}A \) if and only if \( BC = CB \) ), and stated explicitly in [1, 2].

**Theorem 2** Let \( A \in M_n \), \( n \geq 3 \), with \( A\bar{A} = \bar{A}A \). Then there exists a real orthogonal matrix \( Q \in M_n(\mathbb{R}) \) such that \( Q^T AQ \) is of the form:

\[
\Lambda = \begin{bmatrix}
A_1 & * & \cdots & \cdots & * \\
0 & A_2 & * & \cdots & * \\
0 & 0 & A_3 & \cdots & * \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \cdots & A_k
\end{bmatrix},
\]

where each \( A_i \), \( 1 \leq i \leq k \) (for some \( k \)) is a 1-by-1 matrix or a 2-by-2 matrix.

**Example 1.** Note that if \( A = Q\Lambda Q^T \), \( Q \) real orthogonal, \( A\bar{A} = \bar{A}A \) and \( AA^T = A^TA \) if and only if \( \Lambda \bar{A} = \bar{A} \Lambda \) and \( \Lambda^T A = A^T \Lambda \). Also note that if \( A = A^T \) and \( A\bar{A} = \bar{A}A \), then \( A \) is normal since \( A^* = \bar{A} \) in this case.

**Lemma 1** If \( A \in M_2 \) with \( AA^T = A^TA \), then \( A \) is
either symmetric or of the form \[ \begin{bmatrix} a & b \\ -b & a \end{bmatrix}, a, b \in \mathbb{C}. \]

**Proof.** Suppose that \( A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, a, b, c, d \in \mathbb{C} \), with
\[
AA^T = A^T A, \quad \text{then}
\]
\[
\begin{bmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{bmatrix} = \begin{bmatrix} a^2 + c^2 & ab + cd \\ ab + dc & b^2 + d^2 \end{bmatrix}
\]

Hence \( b^2 = c^2 \) and \( ab + cd = ac + bd \).

**Case 1.** \( b = c \), so that \( A \) is symmetric.

**Case 2.** \( b = -c \), then \( ab - bd = ab + bd \) or \( ab = bd \). If \( b = 0 \), then \( A \) is symmetric. If \( b \neq 0 \), \( a = d \) and \( A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \).

**Proposition 1** \( A \in M_2 \) with \( AA^T = A^T A \) and \( AA = \bar{A} \bar{A} \), then \( A \) is normal.

**Proof.** From the Lemma 1, we have two cases. If \( A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}, a, b \in \mathbb{C} \), then \( A \) is normal. On the other hand, if \( A \) is symmetric with \( \bar{A} = \bar{A} \), then since \( A = \bar{A} \) in this case, we must have \( AA = \bar{A} \bar{A} \), so \( A \) is normal.

**Example 2.** We now look at the case of \( A \in M_3 \). We start with a lemma:

**Lemma 2** Suppose \( A \in M_3 \) with \( AA = \bar{A} \bar{A} \), \( AA^T = A^T A \) and \( A = QAQ^T \) for some real orthogonal matrix \( Q \in M_3(\mathbb{R}) \) where \( \Lambda \) is of one of the two forms: see Equation (1).

then \( A \) is normal.

**Proof. Case 1:** \( \Lambda = \begin{bmatrix} a & b & x \\ -b & a & y \\ 0 & 0 & \alpha \end{bmatrix} \). Now we require \( \Lambda \Lambda^T = \Lambda^T \Lambda \), so that
\[
\begin{bmatrix} a & b & x \\ -b & a & y \\ 0 & 0 & \alpha \end{bmatrix} \begin{bmatrix} a & -b & 0 \\ b & a & 0 \\ 0 & 0 & \alpha \end{bmatrix} = \begin{bmatrix} a & b & x \\ -b & a & y \\ 0 & 0 & \alpha \end{bmatrix} \begin{bmatrix} a & -b & 0 \\ b & a & 0 \\ 0 & 0 & \alpha \end{bmatrix}.
\]

Hence \( x^2 + y^2 = 0 \), \( x^2 = z^2 \), \( y^2 + z^2 = 0 \), and \( x = y, x = \pm y, y = \pm z \) and also \( ax = ay \), so \( y = 0 \) (giving \( \Lambda \) diagonal and \( A \) normal) or \( a = c \). Suppose \( y \neq 0 \) so that \( a = c \).

**Case 2(a).** If \( x = -z \neq 0 \), then \( ax = bx + yz \Rightarrow y = b - a \) and \( x = \pm i(b - a) = -z \) so that
\[
\Lambda = \begin{bmatrix} 0 & b & \pm i(b - a) \\ 0 & 0 & \alpha \end{bmatrix}. \quad \text{However, this matrix}
\]
also has the property that \( \bar{\Lambda} = \bar{\Lambda} \bar{\Lambda} \), which gives in Equation (2). It follows from equating the entries in the (1, 2) position
\[
|a|^2 + |b|^2 - 2a\bar{b} = |a|^2 - |b|^2 + 2\bar{a}\bar{b}, \text{ or } |a - b|^2 = 0.
\]

\[
\begin{bmatrix} \alpha^2 + b^2 + x^2 & xy & ax \\ xy & a^2 + b^2 + y^2 & ay \\ ax & ay & x^2 + y^2 + \alpha^2 \end{bmatrix}
\]

It follows that \( x = y = 0 \), and \( A \) is normal.

or
\[
\begin{bmatrix} a^2 + b^2 & 0 & ax - by \\ 0 & a^2 + b^2 & bx + ay \\ ax - by & ay + bx & x^2 + y^2 + \alpha^2 \end{bmatrix}
\]

\[
\begin{bmatrix} a^2 + b^2 & 0 & ax - by \\ 0 & a^2 + b^2 & bx + ay \\ ax - by & ay + bx & x^2 + y^2 + \alpha^2 \end{bmatrix}
\]

\[
\begin{bmatrix} a^2 + b^2 & 0 & ax - by \\ 0 & a^2 + b^2 & bx + ay \\ ax - by & ay + bx & x^2 + y^2 + \alpha^2 \end{bmatrix}
\]

Copyright © 2013 SciRes.
so  \( a = b \), and hence \( \Lambda \) is diagonal and \( A \) is normal.

**Case 2(b).** This is where \( x = z \neq 0 \), and since  
\[ ax = bx + xy \]
we have \( y = a - b \), so that  
\[ \Lambda = \begin{bmatrix} a & \pm i(a-b) & a-b \\ 0 & b & \pm i(a-b) \\ 0 & 0 & a \end{bmatrix}, \]
and this is treated in a similar way.

**Proposition 2** If \( A \in M_3 \) with \( AA^T = A^T A \) and \( AA = A A^T \), then \( A \) is normal.

**Proof.** We show that every case reduces to the case of the Lemma 2. From Theorem 1, every matrix \( \Lambda \) with \( A = Q \Lambda Q^T \) (\( Q \) real orthogonal) can be chosen to be one of the following three forms:

(I) \( \Lambda = \begin{bmatrix} a & x & y \\ 0 & b & z \\ 0 & 0 & c \end{bmatrix} \), (II) \( \Lambda = \begin{bmatrix} a & b & x \\ c & d & y \\ 0 & 0 & e \end{bmatrix} \), or (III) \( \Lambda = \begin{bmatrix} a & b & x \\ 0 & c & y \\ 0 & d & e \end{bmatrix} \).

We have dealt with Case (I) in Lemma 2, so consider Case (II): \( \Lambda \Lambda^T = A^T A \) gives in Equation (3).

It follows that \( b^2 + x^2 = c^2 \), \( c^2 + y^2 = b^2 \), and \( x^2 + y^2 = 0 \), so that \( x = \pm iy \).

**Case 1.** \( x = iy \neq 0 \), then since \( bx + dy = ey \), \( biy + dy = ey \), so \( b = i(d - e) \).

Also \( xe = ax + cy \) gives \( xe = ax - cix \), so that \( c = i(e - a) \), so \( \Lambda \) has the form  
\[ \Lambda = \begin{bmatrix} a & i(d - e) & iy \\ i(e - a) & d & y \\ 0 & 0 & e \end{bmatrix} \]

Now we use the fact that \( \Lambda \Lambda = \Lambda \Lambda^T \). This gives in Equation (4). On equating entries in the (1, 3) position we have:

\[ -a \overline{y} (d - e) + y \overline{e} = a \overline{y} - y (\overline{d} - \overline{e}) - \overline{v} e \]

and simplifying gives  
\[ y (d - a) = y (\overline{d} - \overline{a}) \]

so if \( a \neq d \), we have  
\[ y = -a - d \]

Equating entries in the (2, 3) position gives:

\[ y (e - a) + d \overline{y} + y \overline{e} = y (\overline{e} - \overline{a}) + \overline{d} y + \overline{y} e, \]

and this reduces to:  
\[ y (d - a) = y (\overline{d} - \overline{a}) \]

so if \( a \neq d \),  
\[ y = a - d \]

and we can apply Lemma 2. The other possibility is that \( y = 0 \), so that \( b = \pm c \) and \( \Lambda \) is either of the form  
\[ \Lambda = \begin{bmatrix} a & b & 0 \\ b & d & 0 \\ 0 & 0 & e \end{bmatrix} \]

the form \( \Lambda = \begin{bmatrix} -b & a & y \\ b & a & y \\ 0 & 0 & e \end{bmatrix} \) (when \( c = b \), since in this case \( a = d \)).

**Case 2.** \( x = -iy \neq 0 \), then \( bx + dy = ey \) gives \( b = i(e - d) \), and \( ax + cy = xe \) gives \( c = i(a - e) \), so that \( \Lambda \) has the form  
\[ \Lambda = \begin{bmatrix} a & i(e - d) & iy \\ i(a - e) & d & y \\ 0 & 0 & e \end{bmatrix} \]

We proceed exactly as in Case 1 to reduce \( \Lambda \) to the

\[
\begin{bmatrix}
  a^2 + b^2 + x^2 & ac + bd + xy & xe \\
  ac + bd + xy & c^2 + d^2 + y^2 & ey \\
  xe & ye & e^2
\end{bmatrix}
\]

\[
\begin{bmatrix}
  a^2 + c^2 & ab + cd & ax + cy \\
  ab + cd & b^2 + d^2 & bx + dy \\
  ax + cy & bx + dy & x^2 + y^2 + e^2
\end{bmatrix}
\]

\[
\begin{bmatrix}
  |d|_1 (d - e)(\overline{\alpha} - \overline{a}) - i \alpha (d - e) + i \overline{d} (d - e) - i \alpha \overline{e} + i \overline{y}(d - e) + i y \overline{e} \\
  i \overline{d} (e - a) - i \alpha (e - a)(\overline{d} - \overline{e}) + |d|^2 \overline{y} (e - a) + d \overline{y} + y \overline{e} \\
  0 & 0 & |e|^2
\end{bmatrix}
\]

\[
\begin{bmatrix}
  |d|_1 (\overline{d} - \overline{e})(\overline{\alpha} - \overline{a}) + i \alpha (\overline{d} - \overline{e}) - i \overline{d} (\overline{d} - \overline{e}) + i \overline{y} (\overline{d} - \overline{e}) - i \overline{y} e \\
  -i \alpha (\overline{\alpha} - \overline{a}) + i \overline{d} (e - a)(\overline{\alpha} - \overline{e})(d - e) + |d|^2 \overline{y} + y \overline{e} \\
  0 & 0 & |e|^2
\end{bmatrix}
\]
situation of Lemma 2.

In Case (III), where \( \Lambda = \begin{bmatrix} a & x & y \\ 0 & b & c \\ 0 & d & e \end{bmatrix} \), we proceed exactly as in Case (II) to deduce the result.

REFERENCES

