The intersection of the similarity and conjunctivity equivalence classes

Mark A. Mills

Department of Mathematics and Computer Science, Central College, Pella, IA 50219, USA

Received 12 September 2000; accepted 26 June 2002

Submitted by R.A. Horn

Abstract

Let $A$ be an $n \times n$ complex matrix. Let $\text{Sim}(A)$ denote the similarity equivalence class of $A$, let $\text{Conj}(A)$ denote the conjunctivity equivalence class of $A$, and let $\mathcal{U}(A)$ denote the unitary similarity equivalence class of $A$. Define $\text{CS}(A) = \text{Sim}(A) \cap \text{Conj}(A)$. We seek to classify the matrices that have $\text{CS}(A) = \mathcal{U}(A)$, and show by an example that this is not true in general. But we show that it is true when $A$ is Hermitian or is a scalar multiple of a Hermitian. For $n \geq 3$, we reduce the general $n \times n$ case to the case when $A$ is non-singular and not a multiple of a Hermitian matrix. We completely classify the $2 \times 2$ case and find that $\text{CS}(A) = \mathcal{U}(A)$ when $A$ is non-singular or normal (or both).

© 2002 Elsevier Science Inc. All rights reserved.

Keywords: Similarity; Conjunctivity; Unitary similarity; Equivalence classes

1. Introduction

Let $\text{GL}_n$ denote the set of $n \times n$ complex invertible matrices. Two matrices $A, B \in M_n$ (the set of complex $n \times n$ matrices) are similar if there exists an $S \in \text{GL}_n$ such that $B = S^{-1}AS$. Similarity is an equivalence relation on $M_n$, and we denote the similarity equivalence class of $A \in M_n$ by $\text{Sim}(A)$.

Two matrices $A, B \in M_n$ are conjunctive (or *congruent or Hermitian congruent) if there exists a $T \in \text{GL}_n$ such that $B = T^*AT$, where $T^*$ denotes the conjugate transpose of $T$. Conjunctivity is an equivalence relation on $M_n$, and we denote the conjunctive equivalence class of $A \in M_n$ by $\text{Conj}(A)$.

E-mail address: millsM@central.edu (M.A. Mills).
Two matrices \( A, B \in M_n \) are unitarily similar if there exists \( U \in U_n \) (the set of \( n \times n \) unitary matrices) such that \( B = U^*AU = U^{-1}AU \). Unitary similarity is an equivalence relation on \( M_n \), and we denote the unitary similarity equivalence class of \( A \in M_n \) by \( \mathcal{U}(A) \).

Since \( \mathcal{U}(A) \) is a subset of both \( \text{Sim}(A) \) and \( \text{Conj}(A) \), we know that

\[
\text{CS}(A) \equiv \text{Sim}(A) \cap \text{Conj}(A) \supset \mathcal{U}(A),
\]

so \( \text{CS}(A) \) is non-empty. An example later in this section shows that the containment in (1) can be strict.

For a fixed \( A \in M_n \), we want to determine whether the containment in (1) is actually an equality (i.e., whether \( \text{CS}(A) = \mathcal{U}(A) \)). If not, then there are two or more disjoint unitary similarity classes inside \( \text{CS}(A) \), and we want to determine just how many there are.

We begin by examining Hermitian matrices and scalar multiples of Hermitian matrices (which we call essentially Hermitian matrices). We denote the \( n \times n \) Hermitian matrices by \( \mathcal{H}_n \), and recall that \( A, B \in \mathcal{H}_n \) are unitarily similar if and only if they have the same eigenvalues (counting multiplicities).

**Theorem 1.1.** If \( \alpha A \in \mathcal{H}_n \) for some non-zero \( \alpha \in \mathbb{C} \), then \( \text{CS}(A) = \mathcal{U}(A) \).

**Proof.** Since \( \mathcal{U}(A) \subseteq \text{CS}(A) \), we need to show only the reverse containment. Let \( B \in \text{CS}(A) \) and let \( T, S \in \text{GL}_n \) be such that \( B = T^*AT = S^{-1}AS \). Since \( B = T^*AT \) and \( \alpha A \) is Hermitian, \( \alpha B \) is Hermitian. Since \( B = S^{-1}AS \), \( \alpha A \) and \( \alpha B \) are similar, and so have the same eigenvalues (counting multiplicities). Therefore, since \( \alpha A \) and \( \alpha B \) are unitarily similar, we know that \( A \) and \( B \) are unitarily similar, and \( \text{CS}(A) \subseteq \mathcal{U}(A) \). \( \square \)

The following example shows that Theorem 1.1 is not true for the larger class of normal matrices when \( n \geq 3 \).

**Example 1.2.** For \( n \geq 3 \), consider the \( n \times n \) permutation matrix

\[
P = \begin{bmatrix} 0 & 0 & \ldots & 0 & 1 \\ 1 & 0 & \ldots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \end{bmatrix}.
\]

For any real \( a \neq 0 \), consider the diagonal matrices

\[
D_a = \begin{bmatrix} a & 0 \\ 0 & \frac{1}{a} \end{bmatrix} \oplus I_{n-2}
\]

and

\[
A_a = aI_2 \oplus I_{n-2}.
\]
A computation reveals that
\[ P_a = D_a^* P D_a = \begin{bmatrix}
0 & 0 & 0 & \cdots & 0 & a \\
1 & 0 & 0 & \cdots & 0 & 0 \\
0 & a^{-1} & 0 & \cdots & 0 & 0 \\
0 & 0 & 1 & \cdots & 0 & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 1 & 0
\end{bmatrix} = A_a P A_a^{-1}, \]
so \( P_a \in \text{CS}(P) \). However, the squared Frobenius norms of both \( P \) and \( P_a \) are
\[ \text{tr} \, PP^* = n \]
and
\[ \text{tr} \, P_a^* P_a = n - 2 + a^2 + a^{-2}, \]
respectively. Since the Frobenius norm is invariant under unitary equivalence (and hence under unitary similarity), we conclude that \( P_a \not\in \mathcal{U}(P) \), if \( a \neq 1 \). In fact, this example shows that \( \text{CS}(P) \) contains an uncountable number of disjoint unitary similarity classes, since for \( a > 1 \) we have a continuum of values for \( \text{tr} \, P_a^* P_a \).

For \( 1 \times 1 \) matrices, \( \text{CS}(A) = \mathcal{U}(A) \). In Section 3, we examine the \( 2 \times 2 \) case. However, before we do this, we examine block diagonal matrices.

2. Results about block diagonal matrices

We now examine matrices that can be brought into block diagonal form. We use \( 0_m \) to denote the \( m \times m \) zero matrix.

Before we proceed, we present a theorem that we use to determine when two matrices are unitarily similar. A word in the non-commuting variables \( x \) and \( y \) is a finite formal product of non-negative integer powers of \( x \) and \( y \), and a word’s degree is the sum of all its powers of \( x \) and \( y \).

**Theorem 2.1** [6]. Let \( A, B \in M_n \). Then \( A \) and \( B \) are unitarily similar if and only if \( \text{tr} \, (\omega(A, A^*)) = \text{tr} \, (\omega(B, B^*)) \) for every word \( \omega(x, y) \) of degree less than or equal to \( 2n^2 \). In fact, this suffices to check that \( \text{tr} \, (\omega(A, A^*)) = \text{tr} \, (\omega(B, B^*)) \) for every word \( \omega(x, y) \) of degree less than or equal to \( 2n^2 \). In many cases, we need to check only the word \( \omega(x, y) = xy \); that is, we need to check only that \( A \) and \( B \) have the same Frobenius norm.

**Proposition 2.2.** Let \( A \in M_n \), \( n \geq 2 \). Suppose that \( A \) is conjunctive and similar to
\[ \begin{bmatrix}
B & C \\
0 & 0_{n-r}
\end{bmatrix}, \]
where \( B \in M_r, \ 0 < r < n, \) and \( C \neq 0. \) Then \( \text{CS}(A) \) contains uncountably many disjoint unitary similarity classes.

**Proof.** Without loss of generality, we can assume that 
\[
A = \begin{bmatrix} B & C \\ 0 & 0_{n-r} \end{bmatrix}.
\]
Let \( \alpha \in \mathbb{C} \) with \( \alpha \neq 0, \) and consider 
\[
M = I_r \oplus \alpha I_{n-r}.
\]
Then 
\[
M^*AM = M^{-1}AM = \begin{bmatrix} B & \alpha C \\ 0 & 0_{n-r} \end{bmatrix} \in \text{CS}(A),
\]
\[
\text{tr}((M^*AM)^* (M^*AM)) = \text{tr}(B^*B) + |\alpha|^2 \text{tr}(C^*C)
\]
and for \( \alpha \neq 0 \) we get a continuum of values since \( \text{tr}(C^*C) \neq 0. \) Therefore, Theorem 2.1 shows that \( \text{CS}(A) \) contains uncountably many disjoint unitary similarity classes. \( \Box \)

**Proposition 2.3.** Let \( A, B \in \text{GL}_n, \) and let \( m \) be a given non-negative integer. Then 
\( B \in \text{CS}(A) \) if and only if 
\( B \oplus 0_m \in \text{CS}(A \oplus 0_m). \)

**Proof.** First suppose that \( B \in \text{CS}(A). \) There exist \( T, S \in \text{GL}_n \) such that 
\[
B = T^*AT = S^{-1}AS. \text{ So}
\]
\[
B \oplus 0_m = (T \oplus I_m)^*(A \oplus 0_m)(T \oplus I_m) = (S \oplus I_m)^{-1}(A \oplus 0_m)(S \oplus I_m).
\]
Conversely, let 
\[
T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}
\]
be non-singular block matrices, with \( T_{11}, S_{11} \in M_n, \) such that 
\[
\begin{bmatrix} B & 0 \\ 0 & 0_m \end{bmatrix} = B \oplus 0_m = T^*(A \oplus 0_m)T = \begin{bmatrix} T_{11}^*AT_{11} & T_{11}^*AT_{12} \\ T_{12}^*AT_{11} & T_{12}^*AT_{12} \end{bmatrix}, \quad (3)
\]
and 
\[
S(B \oplus 0_m) = \begin{bmatrix} S_{11}B & 0 \\ S_{21}B & 0_m \end{bmatrix} = \begin{bmatrix} AS_{11} & AS_{12} \\ 0 & 0_m \end{bmatrix} = (A \oplus 0_m)S. \quad (4)
\]
Examination of the \((1,1)\)-entries of \((3)\) shows that \( T_{11}^*AT_{11} = B. \) However, since \( A, B \in \text{GL}_n, \) it follows that \( T_{11} \in \text{GL}_n, \) and so \( A \) and \( B \) are conjunctive. Examination of the \((1,2)\)- and \((2,1)\)-entries of \((4)\) shows that \( AS_{12} = 0 \) and \( S_{21}B = 0. \) Again, since \( A, B \in \text{GL}_n, \) it follows that \( S_{12} = 0 \) and \( S_{21} = 0, \) so \( S \) is block diagonal, \( S_{11}^{-1}AS_{11} = B, \) and \( B \in \text{CS}(A). \) \( \Box \)

**Proposition 2.4** [2, p. 78]. Let \( A, B \in M_n \) and let \( C \in M_p. \) Then \( B \in \Psi(A) \) if and only if \( B \oplus C \in \Psi(A \oplus C). \)
Proof. By Theorem 2.1, \( B \in \mathcal{U}(A) \) if and only if
\[
\text{tr}(\omega(A, A^*)) = \text{tr}(\omega(B, B^*)),
\]
for every word \( \omega(x, y) \) in non-commuting variables \( x \) and \( y \). Since
\[
\omega(X \oplus Y, (X \oplus Y)^*) = \omega(X, X^*) \oplus \omega(Y, Y^*)
\]
for any word \( \omega(x, y) \), (5) is true if and only if
\[
\text{tr}(\omega(A \oplus C, (A \oplus C)^*)) = \text{tr}(\omega(B \oplus C, (B \oplus C)^*))
\]
and by Theorem 2.1, this is true if and only if \( B \oplus C \in \mathcal{U}(A \oplus C) \). \(\square\)

The following theorem shows that for a direct sum of non-singular matrix and a zero matrix, it suffices to focus our attention only on the non-zero block.

**Theorem 2.5.** Let \( A_1 \in \text{GL}_n \), and let \( m \) be a given non-negative integer. Then \( B \in \text{CS}(A_1 \oplus 0_m) \) if and only if \( B \) is unitarily similar to a matrix of the form \( B_1 \oplus 0_m \), for some \( B_1 \in \text{CS}(A_1) \). Furthermore, \( \text{CS}(A_1) = \mathcal{U}(A_1) \) if and only if \( \text{CS}(A_1 \oplus 0_m) = \mathcal{U}(A_1 \oplus 0_m) \).

Proof. If \( B \) is unitarily similar to a matrix of the form \( B_1 \oplus 0_m \), with \( B_1 \in \text{CS}(A_1) \), then clearly \( B \in \text{CS}(A_1 \oplus 0_m) \).

Conversely, suppose \( A_1 \in \text{GL}_n \), let \( A = A_1 \oplus 0_m \), and suppose \( B \in \text{CS}(A) \). Since \( B \) is similar to \( A \), we can apply a unitary similarity and assume that \( B \) is in the block form
\[
\begin{bmatrix}
B_1 & C \\
0 & D
\end{bmatrix},
\]
where \( B_1 \) has the same eigenvalues as \( A_1 \). Since \( A \) and \( B \) have the same rank, we have \( \text{rank}(B) = n \). But \( B_1 \) must be non-singular, because it has the same eigenvalues as \( A_1 \). Hence, \( \text{rank}(B) = n = \text{rank}(B_1) \) and we must have \( D = 0 \). Now let \( B = S^{-1}AS \) for some \( S \in \text{GL}_m \). Theorem 3.5 from [5] then tells us that
\[
S = \begin{bmatrix}
S_{11} & S_{12} \\
0 & S_{22}
\end{bmatrix},
\]
where \( S_{11} \) is \( n \times n \) and non-singular. So \( B_1 = S_{11}^{-1}A_1S_{11} \). We also have \( B = T^*AT \), where \( T \) is non-singular. Let
\[
T = \begin{bmatrix}
T_{11} & T_{12} \\
T_{21} & T_{22}
\end{bmatrix},
\]
with \( T_{11} \in \text{M}_n \). Then
\[
T^*(A_1 \oplus 0_m)T = \begin{bmatrix}
T_{11}^*A_1T_{11} & T_{12}^*A_1T_{12} \\
T_{11}^*A_1T_{11} & T_{12}^*A_1T_{12}
\end{bmatrix} = \begin{bmatrix}
B_1 & C \\
0 & 0_m
\end{bmatrix},
\]
so \( B_1 = T_{11}^*A_1T_{11} \). Since \( A_1 \) and \( B_1 \) are non-singular, \( T_{11} \) must also be non-singular and so \( B_1 \in \text{CS}(A_1) \). We also have \( T_{12}^*A_1T_{11} = 0 \), so that \( T_{12} = 0 \). Hence, \( C = T_{11}^*A_1T_{12} = 0 \) and so \( B \) is unitarily similar to \( B_1 \oplus 0_m \), where \( B_1 \in \text{CS}(A_1) \).
Now suppose that \( \text{CS}(A_1 \oplus 0_m) = \mathcal{U}(A_1 \oplus 0_m) \). Let \( B_1 \in \text{CS}(A_1) \). Then \( B_1 \oplus 0_m \in \mathcal{U}(A_1 \oplus 0_m) \). Let
\[
V = \begin{bmatrix}
V_{11} & V_{12} \\
V_{21} & V_{22}
\end{bmatrix}
\]
be a unitary matrix such that \( B_1 \oplus 0_m = V^*(A_1 \oplus 0_m)V \). Theorem 3.5 of [5] then tells us that \( V_{21} = 0 \). Since \( V \) is unitary, we must then have \( V_{12} = 0 \), so \( V_{11} \) is unitary and \( B_1 \in \mathcal{U}(A_1) \).

Conversely, suppose \( \text{CS}(A_1) = \mathcal{U}(A_1) \), and let \( B \in \text{CS}(A_1 \oplus 0_m) \). Then the first part of the theorem tells us \( B \) is unitarily similar to a matrix of the form \( B_1 \oplus 0_m \), where \( B_1 \in \text{CS}(A_1) \). But then \( B_1 \in \mathcal{U}(A_1) \), and hence \( B \in \mathcal{U}(A_1 \oplus 0_m) \). □

One might hope that the 0
\( m \)
block in Theorem 2.5 can be replaced by any \( m \times m \) matrix (e.g., \( I_m \)) with the result preserved. However, Example 1.2 in the 3 \( \times \) 3 case shows that this is not correct.

The result of our analysis so far is to reduce the problem to examining non-singular matrices that are not essentially Hermitian. Any singular matrix has at least one zero eigenvalue, and so it can be unitarily triangularized with a zero entry in the \((n, n)\) position. If this places the matrix in the form (2), then we know that \( \text{CS}(A) \neq \mathcal{U}(A) \) and we are done. Otherwise, the matrix is a direct sum of a smaller non-singular matrix and a zero matrix, and Proposition 2.4 and Theorem 2.5 say that we need to focus only on the non-singular block matrix in the upper-left corner. If that block is essentially Hermitian, we know how to solve our problem. So we need to consider only the case in which it is not essentially Hermitian.

3. The 2 \( \times \) 2 case

Now that we have reduced the general \( n \times n \) problem to looking at non-singular matrices that are not essentially Hermitian, we stop to look at the 2 \( \times \) 2 case. First recall a theorem that gives a canonical form for 2 \( \times \) 2 matrices under unitary similarity.

**Theorem 3.1** [5, Theorem 2.4]. Let \( A \in M_2 \) have eigenvalues \( \lambda_1 \) and \( \lambda_2 \), and let
\[
r = \sqrt{\text{tr}(A^*A) - |\lambda_1|^2 - |\lambda_2|^2}.
\]
Then \( A \) is unitarily similar to a matrix of the form
\[
\begin{bmatrix}
\lambda_1 & r \\
0 & \lambda_2
\end{bmatrix}
\]
and any upper triangular matrix \( T = [t_{ij}] \) to which \( A \) is unitarily similar has \( |t_{12}| = r \).
This $r$ is commonly called the deflect from normality, and $r = 0$ if and only if $A$ is normal. Also, any normal matrix with eigenvalues that are collinear on a line passing through the origin in $\mathbb{C}$ is essentially Hermitian. In particular, any $2 \times 2$ singular, normal matrix is essentially Hermitian.

We are now set to analyze the $2 \times 2$ case.

**Theorem 3.2.** Let $A \in M_n$. If $A$ is singular and non-normal, then $\text{CS}(A)$ contains uncountably many unitary equivalence classes. Otherwise (i.e., if $A$ is non-singular or normal or both), we have $\text{CS}(A) = \mathcal{U}(A)$.

**Proof.** Theorem 3.1 and Proposition 2.2 take care of the singular and non-normal case. If $A$ is singular and normal, then it is essentially Hermitian and hence $\text{CS}(A) = \mathcal{U}(A)$.

Now assume that $A$ is non-singular and let $B \in \text{CS}(A)$. We must show that $B \in \mathcal{U}(A)$. Applying unitary similarities to triangularize $A$ and $B$, we can assume that

$$A = \begin{bmatrix} \lambda_1 & r \\ 0 & \lambda_2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} \lambda_1 & s \\ 0 & \lambda_2 \end{bmatrix},$$

where $r \geq 0$ and $s \in \mathbb{C}$. Now let $T$ be a non-singular matrix such that $B = T^* AT$. Since $\det(A) = \det(B)$ is non-zero, we have $|\det(T)| = 1$. After multiplying $T$ by a non-zero complex number of modulus one, we can assume that $\det(T) = 1$. Hence, if

$$T = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

then

$$T^{-1} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$  \(\text{(1)}\)

Computing the entries of $BT^{-1} = T^* A$ gives the following four equations:

1. $\lambda_1 d - sc = \lambda_1 \bar{a}$  \(\text{(6)}\)
2. $\lambda_2 d + r\bar{b} = \lambda_2 a$  \(\text{(7)}\)
3. $-\lambda_2 c = \lambda_1 \bar{b}$  \(\text{(8)}\)
4. $-\lambda_1 b + sa = r\bar{a} + \lambda_2 \bar{c}$  \(\text{(9)}\)

Using (6) and (7), we obtain:

5. $cs = \lambda_1 (d - \bar{a})$  \(\text{(10)}\)
6. $\bar{b}r = \lambda_2 (a - \bar{d})$.  \(\text{(11)}\)

If $b = 0$, then (8) tells us that $c = 0$, because $\lambda_2 \neq 0$. Similarly, if $c = 0$, then (8) tells us that $b = 0$. But then (6) (or (7)) says that $d = \bar{a}$, and since $\det(T) = ad = |a|^2 = 1$, we see that $T$ is a diagonal, unitary matrix. So $A$ and $B$ are unitarily similar.
If $b$ and $c$ are both non-zero, then using (10) and (11) leads to
\[ s = \frac{\lambda_1 (d - \overline{a})}{c} \quad \text{and} \quad r = \frac{\lambda_2 (a - \overline{d})}{b}. \]

But (8) then tells us that $\lambda_1 / c = -\lambda_2 / b$, and hence $r = |s|$. So $A$ and $B$ are unitarily similar. \[\Box\]

Now that we have the $2 \times 2$ case completely classified, we can easily find a class matrices that are not essentially Hermitian and for which the conjunctive-similarity equivalence class and unitary similarity class coincide. In particular, if $A_1$ is a $2 \times 2$ non-singular or normal matrix, then $\text{CS}(A_1 \oplus 0_m) = \mathcal{U}(A_1 \oplus 0_m)$.

What happens for non-singular matrices of size $3 \times 3$ or larger that are not essentially Hermitian? If even the normal case could be settled, it would be a step in the right direction. One hindrance is that, as of the writing of this paper, there was no easy way to determine when two matrices are conjunctive. If better criteria for conjunctivity could be produced, they might shed some light on the remainder of the problem. A recent paper by Furtado and Johnson [1] may be useful for making further progress on this problem.

Of course, the intersection of any two matrix equivalence classes can be studied. Some examples include: similarity and $^\mathcal{T}$congruence (complex orthogonal similarity); conjunctivity and consimilarity (complex orthogonal consimilarity); and $^\mathcal{T}$congruence and consimilarity (unitary consimilarity). My Ph.D. dissertation [3] also includes some analysis of the intersection of the similarity equivalence class and the unitary equivalence equivalence class (matrices with the same singular values).

Acknowledgements

I would like to express my gratitude to Professor Bryan E. Cain for his guidance during the researching and writing of my Ph.D. dissertation, which includes the results in this paper. I am also very thankful to the editor and referees for their thoughtful comments and suggestions.

References