PSEUDO-SIMILARITY FOR MATRICES OVER A FIELD

R. E. HARTWIG AND F. J. HALL

ABSTRACT. We call two square matrices $A$ and $B$ (over a ring) pseudo-similar, if matrices $X$, $X^{-}$, $X^{\sim}$ exist, such that $X^{-}AX = B$, $XBX^{\sim}A$, $XX^{\sim}X = X$ and $XX^{\sim}X = X$. We show that if $A$ and $B$ have the same dimension and if the ring is a field, then pseudo-similarity implies similarity, and hence that pseudo-similarity is an equivalence relation.

1. Introduction and notation. In this note we shall investigate the algebraic properties of pseudo-similarity which we define as follows.

DEFINITION. Let $\mathbb{R}_{m \times n}$ denote the $m \times n$ matrices over a ring with unity. If $A \in \mathbb{R}_{m \times m}$ and $B \in \mathbb{R}_{n \times n}$, we say that $A$ is pseudo-similar to $B$, via $X$ and we write $A \sim B$, if there exists $X \in \mathbb{R}_{m \times n}$ and two possibly distinct $X^{-}$, $X^{\sim} \in \mathbb{R}_{m \times m}$ such that

\begin{align*}
(1) & \quad X^{-}AX = B, \\
(2) & \quad XBX^{\sim} = A, \\
(3) & \quad XX^{-}X = X, \\
(4) & \quad XX^{\sim}X = X.
\end{align*}

In general pseudo-similarity does not imply similarity as seen from the following example of $1 \times 1$ matrices over the ring $\mathcal{L}(\mathbb{R}^{\infty})$, of linear transformations on the vector space of real sequences.

Let $X$ be the right shift and let $Y$ be the left shift. Suppose furthermore that $A = XY$ and $B = I$. Then $XYX = X$, $YAX = B$ and $XBY = A$, so that $A \sim B \not\sim A$.

The theorem in the abstract is established by use of the core-nilpotent decomposition of a square matrix over a field. First we give algebraic properties of pseudo-similarity in a general setting.

As always, any solution to $XMX = X$ will be denoted by $X^{-}$, $X^{\sim}$ etc., while $X^{+}$ denotes any solution to $XMX = X$ and $MXM = M$. These generalized inverses are sometimes called “inner” and “reflexive” inverses of $X$, respectively [1]. The set of all inner (reflexive) inverses of $(\cdot)$ will be indicated by $(\cdot)^{-}$, $(\cdot)^{+}$. Ranges, row spaces and nullspaces of $(\cdot)$ will be denoted by $R(\cdot)$, $RS(\cdot)$ and $N(\cdot)$, respectively.

Received by the editors February 20, 1976.


Key words and phrases. Pseudo-similarity, inner inverse, matrix equations.

© American Mathematical Society 1978
2. Properties of \( \sim \).

**Theorem 1.** Suppose \( A \sim B \) with \( X, X^- \) as in the definition (a). Then the following are valid.

1. \( A = XX^-AXX^- = AXX^- = XX^-A, \quad B = X^-XBX^-X = X^-XB = BX^-X \).

2. \( X^-A^kX = B^k, \quad XB^kX^- = A^k, \quad k = 1, 2, \ldots \).

3. \( \{A^-X\} \subseteq \{(X^-A)\}, \quad \{XB^-\} \subseteq \{(BX^-)\}, \quad \{A^+X\} \subseteq \{(X^-A)^+\}, \quad \{XB^+\} \subseteq \{(BX^-)^+\} \).

4. (i) \( \{X^-A^-X\} \subseteq \{B^-\}, \quad \{XB^-X^-\} \subseteq \{A^-\}, \quad \{X^-A^+X\} \subseteq \{B^+\} \), (iv) \( \{XB^+X^-\} \subseteq \{A^+\} \).

5. (i) \( \{X^-AXB^-X^-X\} \subseteq \{B^-\} \), (ii) \( \{XX^-A^-XX^-\} \subseteq \{A^-\} \), (iii) \( \{X^+A^-X\} \subseteq \{B^+\} \).

6. \( \{X^-AA^+X^-\} \subseteq \{(BB^-)\}^+, \quad \{X^+A^-AX\} \subseteq \{(B^-B)\}^+, \quad \{XBB^-X^-\} \subseteq \{(AA^-)\}^+, \quad \{XB^-BX^-\} \subseteq \{(A^-A)\}^+ \).

7. (i) \( B^-B = I \Rightarrow X^-X = I \Leftrightarrow XX^-X = I \Leftrightarrow BB^- = I \), (ii) \( A^-A = I \Rightarrow XX^-X = I \Leftrightarrow AA^- = I \).

8. \( R(A) \cong R(B), \quad RS(A) \cong RS(B) \) as \( \mathcal{R} \)-modules.

In these expressions \( X^- \) and \( X^+ \) are the inner inverses of \( X \) that occur in (1)–(4), while \( A^- \), \( B^- \) are any inner inverses of \( A, B \).

**Proof.** (1) Substitution of \( X^-AX = B \) in \( XBX^- = A \) yields \( A = XX^-AXX^- \), which on successive premultiplication by \( XX^- \) and post-multiplication by \( XX^- \) shows that \( A = XX^-A = AXX^- \). Similarly \( B = X^-XBX^-X = X^-XB = BX^-X \).

(2) Clearly \( B^2 = X^-A(XX^-A)X = X^-A^2X \) and \( A^2 = X(BX^-X)BX^- = XB^2X^- \). The proof now follows by induction.


The remaining results follow by symmetry.


(5) These follow from (4) on substituting 4(ii) in 4(i) and vice-versa. Similarly for 4(iii) and 4(iv).

\[
\]
One can establish the other results in an analogous way.

(7) From (6), \( BB^- = I \Rightarrow X^-AA^-X = I \) and hence \( X^-X = X^-AA^-XX^-X = X^-AA^-X = I \). Similarly for the other results.

(8) Define \( \varphi: R(A) \to R(B) \) by \( \varphi(As) = BX^-s, \quad s \in \mathbb{R}^n \). Then \( \varphi \) is clearly a well-defined \( \mathbb{R} \)-homomorphism. If \( BX^-s_1 = BX^-s_2 \) then \( As_1 = XBX^-s_1 = XBX^-s_2 = As_2 \), so that \( \varphi \) is one-one.
Lastly, \( By = BX^\sim Xy = \varphi (AXy) \), implying \( \varphi \) is also onto. Thus \( R(A) \cong R(B) \). In the same way, if we define \( \theta (u^T A) = u^T XB \), it then follows that \( RS(A) \cong RS(B) \). The proof of the theorem is now complete.

Let us now rephrase the condition of pseudo-similarity in a form which is more transparent when viewed from the theory of matrix equations.

**Corollary 1.** The following are equivalent.

1. \( A \sim B \) via \( X \).
2. \( AX = XB \), \( AXAX = BXX = X = X \).
3. \( BX = X = AX = A, BX = BXX = XX = X \).

**Proof.** (i) \( \Leftrightarrow \) (ii). Let \( A \sim B \). By Theorem 1, part 1, \( A = AXX = XAX = A \), \( B = XBX = XBX = B \). Hence \( XAX = B \Rightarrow XB = XXAX = AX \).

Conversely substituting \( AX = XB \) in \( A = AXX = XAX \) and \( B = XBX = XBX \) yields \( A = XBX \) and \( B = XBX \) respectively. The equivalence of (i) and (iii) follows similarly.

**Remarks.**

1. One can use these conditions to verify the following range and row-space equations: \( R(A) = R(XB), R(B) = R(XAX), RS(A) = RS(BX), RS(B) = RS(AX) \).

2. The equivalence between (ii) and (iii) also holds without \( X, X \) being inner inverses.

**3. Main Theorem.** Let us now specialize to the case where \( R \) is a field \( \mathbb{F} \), and suppose that \( A, B \) and \( X \) have the same dimension \( n \). In this case the product rule for determinants shows that the main assertion (Theorem 2) is obvious if any of \( A, B \) or \( X \) is invertible.

Hence without loss of generality we may assume all three matrices to be singular.

The following additional corollaries of Theorem 1 are then immediate.

1. If \( p(\lambda) = p_0 + p_1 \lambda + \cdots + p_n \lambda^n \) and \( p(A) = 0 \), then the minimal polynomial \( \psi_A \) divides \( p(\lambda) \) and hence \( p_0 = 0 \).

Thus by part 2 of Theorem 1,

\[
0 = \sum_{i=1}^{n} p_i A^i \Rightarrow 0 = \sum_{i=1}^{n} p_i XAX = \sum_{i=1}^{n} p_i B^i = p(B).
\]

The converse follows by symmetry. In particular, the minimal polynomials \( \psi_A \) and \( \psi_B \) are equal, so that \( A \) and \( B \) have equal indices, which are defined by index(\( A \)) = \( \min_k \{ R(A^k) = R(A^{k+1}) \} \) [1, p. 170]. Also, if \( \Delta_M \) denotes the characteristic polynomial of \( M \), then \( \Delta_B = \Delta_X = \Delta_{XX} = \Delta_A \), since \( \Delta_{PQ} = \Delta_{QP} \) always holds for square matrices \( P \) and \( Q \) over a commutative ring.

(2) The ranks of \( A^k \) and \( B^k \) must be equal for all \( k = 0, 1, \ldots \), as seen from part 2. This has far reaching consequences. Indeed, this implies that \( A \) and \( B \) have the same nilpotent Jordan blocks \( J_i(0) \), because the zero Weyr characteristics are the same [2]. This may also be seen from Fitting’s core-nilpotent decomposition \( P^{-1} AP = \text{diag} [U_A, N_A] \), where \( U_A \) is invertible and
$N_A$ contains all the nilpotent Jordan blocks (even over nonclosed fields). Thus $A(I - AA^d) \approx B(I - BB^d)$, where $(\cdot)^d$ is the Drazin inverse of $A$, defined by $P^{-1}AP = \text{diag}[U_A^{-1}, 0]$ [1, p. 169].

**Theorem 2.** Let $A, B \in \mathbb{F}_{n \times n}$. Then $A \approx B \iff A \approx B$.

**Proof.** The sufficiency is clear. Suppose therefore that $XAX = B$ and $XBX = A$ for some $X^{-1}, X \in \mathbb{F}_{n \times n}$, so that the conditions of Corollary 1, part (ii), hold. As observed earlier, we may without loss of generality assume that $A, B$ and $X$ are singular. Now, consider the core-nilpotent decomposition: [1, p. 175],

$$P^{-1}AP = JA = \text{diag}[U_A, N_A], \quad Q^{-1}BQ = JB = \text{diag}[U_B, N_B].$$

Then $AX = XB, AXX = A, XXB = B$ imply $J_A Y = YJ_B, J_A YY = J_A, Y^{-1}YJB = JB$ where $Y = PXXQ, Y^{-1} = Q^{-1}XXP$. Since $A \approx B, \Delta_A = \Delta_B$ as noted above and it follows that $U_A$ and $U_B$ have the same size. In addition, Theorem 1, part 2, ensures that $N_A = N_B$. We now proceed by partitioning $Y$ conformally as

$$Y = \begin{bmatrix} Y_1 & Y_3 \\ Y_2 & Y_4 \end{bmatrix}$$

and substitute in $J_A Y = YJ_B$. This yields

$$U_A Y_3 = Y_3 N_B, \quad N_A Y_2 = Y_2 U_B, \quad U_A Y_1 = Y_1 U_B, \quad N_A Y_4 = Y_4 N_B. \quad (3.1)$$

Since there are no common elementary divisors for $U_A$ and $N_B$ it follows that $[3] Y_3 = 0$. Similarly $Y_2 = 0$. Suppose we now partition $Y$ as

$$Y^{-1} = \begin{bmatrix} S & * \\ * & * \end{bmatrix},$$

and hence

$$Y^{-1}Y = \begin{bmatrix} SY_1 & * \\ * & * \end{bmatrix}.$$  

So, $Y^{-1}YJB = JB$ shows in particular that $SY_1 U_B = U_B$. Since $U_B$ is invertible we may deduce that $Y_1$ is also invertible. Returning to (3.1) we then have $Y_1^{-1}U_A Y_1 = U_B$ and hence that

$$\begin{bmatrix} Y_1^{-1} & 0 \\ 0 & I \end{bmatrix} J_A \begin{bmatrix} Y_1 & 0 \\ 0 & I \end{bmatrix} = J_B,$$

or equivalently, $A \approx B$, as desired.

If $A$ and $B$ are square but of different dimensions then we may add zeros to the matrices $X, X^{-1}$ etc., to obtain the following generalization of the above result.

**Corollary 2.** If $A \in \mathbb{F}_{m \times m}, B \in \mathbb{F}_{n \times n}$ and $m > n$, then $A \approx B \iff A \approx B \iff A \approx \text{diag}[B, 0]$.

**Remarks.** What we have proven is in fact that pseudo-similarity is an
equivalence relation. In particular, if \( X^{-}AX = B \), \( XBX^{-} = A \), \( Y^{-}BY = C \) and \( YCY^{-} = B \), then there exist \( Z \), \( Z^{-} \) and \( Z = \) such that \( Z^{-}AZ = C \) and \( ZCZ^{-} = A \). This matrix \( Z \) however, need not equal \( XY \), nor \( Z^{-} \) equal \( Y^{-}X^{-} \), even though \[
(Y^{-}X^{-})A(YY) = C \quad \text{and} \quad (YY)C(Y^{-}X^{-}) = A.
\]

This relation does hold if \( Y^{-}X^{-} \) and \( Y^{-}X^{-} \) are both inner inverses of \( XY \). Equivalent local conditions for this were obtained in [4, Proposition 2.1(3)]. It may be shown, using the general form \( \{ Y^{-}\} = Y^{-} + (I - Y^{-}Y)H + K(I - YY^{-}) \), \( H, K \) arbitrary, that in fact \( \{ Y^{-}\}\{X^{-}\} \subseteq \{(XY)^{-}\} \Rightarrow N(X) \subseteq R(YY) \), which clearly will not (and cannot) be true in general.

We close with the conjecture that Theorem 2 holds for matrices over a unit-regular ring, for which each \( a \in R \) has at least one unit (= invertible) inner inverse \( a^{-} \).

**REFERENCES**


**DEPARTMENT OF MATHEMATICS, NORTH CAROLINA STATE UNIVERSITY, RALEIGH, NORTH CAROLINA 27607**

**DEPARTMENT OF MATHEMATICS, PEMBROKE STATE UNIVERSITY, PEMBROKE, NORTH CAROLINA 28372** (Current address of F. J. Hall)

**Current address** (R. E. Hartwig): Department of Mathematics, University of Graz, Graz, Austria