PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 71, Number 1, August 1978

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^{-}, $X^{=}$ exist, such that $X^{-}AX = B$, $XBX^{=}A$, $XX^{-}X = X$ and $XX^{=}X = X$. 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 $\mathfrak{R}_{m\times n}$ denote the $m \times n$ matrices over a ring with unity. If $A \in \mathfrak{R}_{m\times m}$ and $B \in \mathfrak{R}_{n\times n}$, we say that A is pseudo-similar to B, via X and we write $A \approx B$, if there exists $X \in \mathfrak{R}_{m\times n}$ and two possibly distinct X^{-}, $X^{=} \in \mathfrak{R}_{n\times m}$ such that

$$\begin{aligned} &(1) \quad X^{-}AX = B,\\ &(2) \quad XBX^{=} = A,\\ &(3) \quad XX^{-}X = X,\\ &(4) \quad XX^{=}X = X. \end{aligned} \qquad (\alpha)$$

In general pseudo-similarity does not imply similarity as seen from the following example of 1×1 matrices over the ring $\mathfrak{L}(\mathbf{R}^{\infty})$, 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 $A \approx B \not\approx A$.

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^{-}, $X^{=}$ 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.
AMS (MOS) subject classifications (1970). Primary 15A09, 15A24.
Key words and phrases. Pseudo-similarity, inner inverse, matrix equations.
© American Mathematical Society 1978

2. Properties of $\approx$.

THEOREM 1. *Suppose* $A \approx B$ *with* X, X^-, $X^=$ *as in the definition* (α). *Then the following are valid.*

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

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

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

(4) (i) $\{X^=A^-X\} \subseteq \{B^-\}$, (ii) $\{XB^-X^-\} \subseteq \{A^-\}$, (iii) $\{X^=A^+X\} \subseteq \{B^+\}$, (iv) $\{XB^+X^-\} \subseteq \{A^+\}$.

(5) (i) $\{X^=XB^-X^-X\} \subseteq \{B^-\}$, (ii) $\{XX^=A^-XX^-\} \subseteq \{A^-\}$, (iii) $\{X^=XB^+X^-X\} \subseteq \{B^+\}$, (iv) $\{XX^=A^+XX^-\} \subseteq \{A^+\}$.

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

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

(8) $R(A) \cong R(B)$, $RS(A) \cong RS(B)$ *as* $\mathfrak{R}$-*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 postmultiplication 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.

(3) Obviously, $X^-A(A^-X)X^-A = X^-AA^-(XX^-A) = X^-AA^-A = X^-A$ and $A^+X(X^-A)A^+X = A^+(XX^-A)A^+X = A^+AA^+X = A^+X$.

The remaining results follow by symmetry.

(4) We verify that $B(X^=A^-X)B = (BX^=)A^-(XB) = X^-AA^-AX = X^-AX = B$ and $(X^=A^+X)B(X^=A^+X) = X^=A^+(XBX^=)A^+X = X^=A^+AA^+X = X^=A^+X$. The remaining parts 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).

(6) First note that $BB^-(X^-AA^-X)BB^- = BB^-X^-AA^-(XB)B^- = BB^-X^-AA^-AXB^- = BB^-BB^- = BB^-$. Secondly,

$$(X^-AA^-X)BB^-(X^-AA^-X) = X^-AA^-XBB^-(X^-A)A^-X$$
$$= X^-AA^-XBB^-BX^=A^-X = X^-AA^-AA^-X = X^-AA^-X.$$

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 φ: $R(A) \to R(B)$ by $\varphi(A\mathbf{s}) = BX^=\mathbf{s}$, $\mathbf{s} \in \mathfrak{R}^n$. Then φ is clearly a well-defined $\mathfrak{R}$-homomorphism. If $BX^=\mathbf{s}_1 = BX^=\mathbf{s}_2$ then $A\mathbf{s}_1 = XBX^=\mathbf{s}_1 = XBX^=\mathbf{s}_2 = A\mathbf{s}_2$, so that φ is one-one.

Lastly, $By = BX^{=}Xy = \varphi(AXy)$, implying φ is also onto. Thus $R(A) \cong R(B)$. In the same way, if we define $\theta(\mathbf{u}^T A) = \mathbf{u}^T XB$, 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.*

(i) $A \approx B$ *via* X.

(ii) $AX = XB, AXX^{=} = A, X^{-}XB = B, XX^{-}X = XX^{=}X = X$.

(iii) $BX^{=} = X^{-}A, XX^{-}A = A, BX^{=}X = B, XX^{-}X = XX^{=}X = X$.

PROOF. (i) $\Leftrightarrow$ (ii). Let $A \approx B$. By Theorem 1, part 1, $A = AXX^{=} = XX^{-}A$, $B = X^{-}XB = BX^{=}X$. Hence $X^{-}AX = B \Rightarrow XB = XX^{-}AX = AX$. Conversely substituting $AX = XB$ in $A = AXX^{=}$ and $B = X^{-}XB$ yields $A = XBX^{=}$ and $B = X^{-}AX$ 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: $R(A) = R(XB)$, $R(B) = R(X^{-}A)$, $RS(A) = RS(BX^{=})$, $RS(B) = RS(AX)$.

(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 $\mathfrak{R}$ is a field $\mathfrak{F}$, 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 ψ_A divides $p(\lambda)$ and hence $p_0 = 0$.

Thus by part 2 of Theorem 1,

$$0 = \sum_1^n p_i A^i \Rightarrow 0 = \sum_1^n p_i X^{-} A^i X = \sum_1^n p_i B^i = p(B).$$

The converse follows by symmetry. In particular, the minimal polynomials ψ_A and ψ_B are equal, so that A and B have equal indices, which are defined by $\text{index}(A) = \min_k\{R(A^k) = R(A^{k+1})\}$ [**1**, p. 170]. Also, if Δ_M denotes the characteristic polynomial of M, then $\Delta_B = \Delta_{X^{-}AX} = \Delta_{XX^{-}A} = \Delta_A$, 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, . . . , 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 $P^{-1}A^dP = \text{diag}[U_A^{-1}, 0]$ [1, p. 169].

THEOREM 2. *Let $A, B \in \mathfrak{F}_{n\times n}$. Then $A \approx B \Leftrightarrow A \approx B$.*

PROOF. The sufficiency is clear. Suppose therefore that $X^{-}AX = B$ and $XBX^{=} = A$ for some $X^{-}, X^{=} \in \mathfrak{F}_{n\times n}$, 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],

$$P^{-1}AP = J_A = \text{diag}[U_A, N_A], \qquad Q^{-1}BQ = J_B = \text{diag}[U_B, N_B].$$

Then $AX = XB$, $AXX^{=} = A$, $X^{-}XB = B$ imply $J_AY = YJ_B$, $J_AYY^{=} = J_A$, $Y^{-}YJ_B = J_B$ where $Y = P^{-1}XQ$, $Y^{-} = Q^{-1}X^{-}P$ and $Y^{=} = Q^{-1}X^{=}P$.

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 $J_AY = YJ_B$. This yields

$$U_AY_3 = Y_3N_B, \quad N_AY_2 = Y_2U_B, \quad U_AY_1 = Y_1U_B, \quad N_AY_4 = Y_4N_B. \qquad (3.1)$$

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

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

and hence

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

So, $Y^{-}YJ_B = J_B$ shows in particular that $SY_1U_B = U_B$. Since U_B is invertible we may deduce that Y_1 is also invertible. Returning to (3.1) we then have $Y_1^{-1}U_AY_1 = U_B$ 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^{-} etc., to obtain the following generalization of the above result.

COROLLARY 2. *If $A \in \mathfrak{F}_{m\times m}$, $B \in \mathfrak{F}_{n\times n}$ and $m > n$, then $A \approx B \Leftrightarrow A \approx$ 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

$$(Y^-X^-)A(XY) = C \quad \text{and} \quad (XY)C(Y^=X^=) = A.$$

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 $\{Y^-\}\{X^-\} \subseteq \{(XY)^-\} \Leftrightarrow N(X) \subseteq R(Y)$, 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

1. A. Ben-Israel and T. N. E. Greville, *Generalized inverses*, Theory and Applications, Wiley, New York, 1974.

2. R. E. Hartwig, *Some properties of hypercompanion matrices*, J. Industrial Math. **24** (1974), 77–84.

3. ______, *The resultant and the matrix equation $AX = XB$*, SIAM J. Appl. Math. **22** (1972), 538–544.

4. N. Shinozaki and M. Sibuya, *The reverse order law $(AB)^- = B^-A^-$*, Linear Algebra and Appl. **9** (1974), 29–40.

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