SIAM J. APPL. MATH.
Vol. 24, No. 4, June 1973

# THE PSEUDO-INVERSE OF A PRODUCT*

RICHARD BOULDIN†

**Abstract.** Let $A$ and $B$ be bounded linear operators on a complex Hilbert space $H$, such that the range of each is a closed subspace of $H$. The following three conditions are necessary and sufficient for the pseudo-inverse of $AB$ to be the pseudo-inverse of $A$ followed by the pseudo-inverse of $B$: (i) the range of $AB$ must be closed; (ii) the range of $A^*$ must be invariant under $BB^*$; (iii) the intersection of the range of $A^*$ and the kernel of $B^*$ must be invariant under $A^*A$. We use this basic result to obtain a simple technique for computing the pseudo-inverse of a given operator, particularly a given matrix.

**1. Introduction.** The notion of the pseudo-inverse is a concept frequently used in operator theory which has some practical concrete applications. See [8], [9], and the attached bibliographies for material on the applications. A basic problem in the operator theory of the pseudo-inverse is to determine when the pseudo-inverse of a product is the product of the pseudo-inverses. In this paper we give three conditions which are both necessary and sufficient for the pseudo-inverse of $AB$ to be the product of the pseudo-inverse of $B$ and that of $A$. We then apply this result to the problem of computing the pseudo-inverse of a given operator by applying our main result to a factorization of the given operator. Although we consider the polar factorization which is a standard tool in operator theory, we give another very simple factorization which is particularly suited to the problem of determining the pseudo-inverse of a given operator. The results of [6] show that there is a different factorization of the given operator $T$ for every operator $B$ with the property that the range of $B$ contains the range of $T$.

The book [4] which appeared subsequent to the initial preparation of this manuscript has some results for finite-dimensional matrices similar to our results for arbitrary operators; see [4, Thm. 2, p. 13].

**2. Preliminaries.** By "operator" we mean a bounded linear transformation of the complex Hilbert space $H$ into itself. We shall denote the kernel or null space of an operator $T$ by ker $T$ and the orthogonal complement of a subspace $M$ is denoted $M^{\perp}$. If $M$ is a subspace of $H$ invariant under $T$, then $T/M$ denotes the restriction of $T$ to $M$. We shall use, many times in this paper, the elementary fact that $(T^*H)^{\perp}$ is ker $T$ for any operator $T$.

We now recall the definition of the pseudo-inverse. If $T$ is restricted to the orthogonal complement of its kernel, then it defines a one-to-one transformation into the Hilbert space $TH^-$ (the bar means topological closure). Define the pseudo-inverse of $T$, denoted $T^+$, to be the operator on $H$ which is zero on $(TH)^{\perp}$, and on $TH$ it is the inverse to the above transformation induced by $T$. In the following theorem we prove several elementary facts about the pseudo-inverse.

THEOREM 2.1. (i) *The range of $T$ is closed if and only if $T^+$ is bounded.*

(ii) *For any operator $T$ the product $T^+T$ is the orthogonal projection onto $T^*H^-$.*

---

* Received by the editors November 11, 1971, and in revised form September 22, 1972.

† Department of Mathematics, University of Georgia, Athens, Georgia 30601.

(iii) *Provided that $T$ has closed range, the product $TT^+$ is the orthogonal projection onto $TH$.*

*Proof of* (i). The definition of $T^+$ as the inverse of a certain closed linear transformation shows that $T^+$ is a closed linear transformation. If $TH$ is closed, then the closed graph theorem implies that $T^+$ is bounded. On the other hand, assume that $T^+$ is bounded and let $\{Tf_n\}$ be a convergent sequence. With no loss of generality we may assume that $\{f_n\}$ is orthogonal to the kernel of $T$; the continuity of $T^+$ implies that $\{f_n = T^+Tf_n\}$ converges. Clearly, $T$ carries the limit of $\{f_n\}$ onto the limit of $\{Tf_n\}$ and so $TH$ is closed.

*Proof of* (ii) *and* (iii). The linear transformation $T^+T$ is defined on every vector and it is the identity on the orthogonal complement of ker $T$. Since it is zero on ker $T$, it coincides with the orthogonal projection onto the orthogonal complement of ker $T$, which is the closure of $T^*H$. The last assertion of the theorem is proved similarly.

Even if $T$ does not have closed range, the pseudo-inverse of $T$ exists as a closed densely defined linear transformation.

The next result is a key lemma in the proof of our main theorem.

LEMMA 2.2. *Let $P$ and $Q$ be the orthogonal projections onto the subspaces $M$ and $N$, respectively. The restriction $PQ/M$ is the identity on $M$ if and only if $N$ contains $M$.*

*Proof.* This is immediate from [10, Thm. 2, p. 48].

In [3] we introduced the notion of angle between two subspaces and we proved some results that we shall use in this paper. If $M$ and $N$ are subspaces of $H$ then we define the angle between $M$ and $N$ by analogy to finite-dimensional Euclidean space. The angle is between 0 and $\pi/2$ and its cosine is the following supremum:

$$\sup\{|\langle f, g\rangle| : f \in M, g \in N \text{ and } \|f\| = 1 = \|g\|\}.$$

If either $M$ or $N$ is trivial, then the angle is $\pi/2$.

We conclude the preliminaries by noting the following straightforward facts which will be used in the proof of our main theorem.

LEMMA 2.3. *Assume that each of the operators $A$, $B$ and $AB$ has a bounded pseudo-inverse. If $H_0 =$ (ker $AB)^\perp$, then $BH_0$ is a closed subspace of $H$. Furthermore, if $P$ is the orthogonal projection onto* (ker $A)^\perp$, *then $PBH_0$ is a closed subspace of $H$.*

*Proof.* This lemma is based on the well-known fact that $TH$ is closed if and only if $\inf\{\|Tf\| : f \in (\ker T)^\perp, \|f\| = 1\}$ is positive; see [11, p. 231].

Since ker $AB$ contains ker $B$, we know that $H_0$ is contained in (ker $B)^\perp$. If $BH_0$ were not closed there would be a sequence of unit vectors in $H_0$, say $\{f_n\}$, such that $\{\|Bf_n\|\}$ would converge to zero. Since such a sequence would be contained in (ker $B)^\perp$, its existence would contradict that $BH$ is closed. The last sentence above has a similar proof which leads to a contradiction of the fact that $ABH_0$ is closed.

## 3. Main results.

THEOREM 3.1. *Let $A$ and $B$ be operators with bounded pseudo-inverses. For $(AB)^+$ to be bounded it is necessary and sufficient that the angle between $BH$ and*

$\ker A \cap (\ker A \cap BH)^{\perp}$ *be positive. Furthermore, the equation*

$$(AB)^{+} = B^{+}A^{+} \tag{1}$$

*holds if and only if the following three conditions hold*:

$$(AB)^{+} \text{ is bounded}, \tag{2}$$

$$A^{*}H \text{ is invariant under } BB^{*}, \tag{3}$$

$$A^{*}H \cap \ker B^{*} \text{ is invariant under } A^{*}A. \tag{4}$$

*Proof.* By Theorem 2.1, both $A$ and $B$ have closed range, and by the main theorem of [3] the angle condition is equivalent to $(AB)$ having closed range. Again by Theorem 2.1, $(AB)$ having closed range is equivalent to $(AB)^{+}$ being bounded. Thus it suffices to prove the last sentence of the theorem.

First we shall assume that (1) holds and we shall deduce (2), (3) and (4). Because $B^{+}$ and $A^{+}$ are both bounded, we see that (2) holds. Let $H_0 = (\ker AB)^{\perp}$ and define $B'$ to be the restriction of $B$ to $H_0$ considered as a transformation into $BH_0$. Let $P'$ be the transformation from $BH_0$ into $PBH_0$, where $P$ is the orthogonal projection onto $A^{*}H = (\ker A)^{\perp}$ and $P'$ is the restriction of $P$ to $BH_0$. Finally let $A'$ be the transformation from $PBH_0$ into $ABH_0$ defined by restricting $A$ to $PBH_0$ while noting that

$$AB = A[P + (I - P)]B = APB + A(I - P)B = APB. \tag{5}$$

Clearly, for any $f \in H_0$ we have

$$A'P'B'f = ABf \tag{6}$$

and since $AB$ restricted to $H_0$ is one-to-one, it must be that each of the transformations $A'$, $P'$, and $B'$ is one-to-one. Clearly each of these transformations is onto by construction and thus

$$(A'P'B')^{-1} = B'^{-1}P'^{-1}A'^{-1}.$$

In view of (6) and the definition of the pseudo-inverse, for $g \in ABH$ we have

$$(AB)^{+}g = (A'P'B')^{-1}g = B'^{-1}P'^{-1}A'^{-1}g. \tag{7}$$

A moment of reflection on the definition of $A'$ and $B'$ will show that (7) implies that

$$(AB)^{+}g = B^{+}P'^{-1}A^{+}g. \tag{8}$$

Recall that Theorem 2.1 showed $BB^{+}$ to be the orthogonal projection onto $BH$ which contains $BH_0$. We substitute from (1) into (8) and then we apply $B$ to both sides of the resulting equation

$$BB^{+}A^{+}g = BB^{+}P'^{-1}A^{+}g$$

or

$$QA^{+}g = P'^{-1}A^{+}g, \tag{9}$$

where $Q$ is the orthogonal projection onto $BH$. Since $g$ belongs to $ABH_0$, we have $g = ABf$ with $f \in H_0$ and this substituted in (9) results in equation

$$QPBf = P'^{-1}PBf = P'^{-1}P'Bf = Bf. \tag{10}$$

Let $Q_0$ be the orthogonal projection onto $BH_0$ and note that $Q_0Q = Q_0$; apply $Q_0$ to both sides of (10) to get

$$Q_0PBg = Bf.$$

Since this holds for any $f \in H_0$ we invoke Lemma 2.2 to conclude that $BH_0$ is contained in $A^*H$. Since $(\ker AB)^\perp = (AB)^*H = B^*A^*H$, we have

$$A^*H \supset BB^*A^*H.$$

This condition is clearly equivalent to the invariance of $A^*H$ under $BB^*$, so (3) holds.

Now we shall show that (4) follows from (1). Since $A^*H$ is clearly invariant under $A^*A$, we have

$$A^*A(A^*H \cap \ker B^*) \subset A^*H$$

and it suffices to show that

$$A^*A(A^*H \cap \ker B^*) \subset \ker B^* = (BH)^\perp. \tag{11}$$

This is equivalent to showing that for each $f \in A^*H \cap \ker B^*$ and each $g \in H$ we have

$$\langle A^*Af, Bg \rangle = 0, \quad \text{or} \quad \langle Af, ABg \rangle.$$

Thus (11) is equivalent to the equation

$$A(A^*H \cap \ker B^*) \subset (ABH)^\perp. \tag{12}$$

Since (1) holds, we know that the kernel of $B^+A^+$ is exactly $(ABH)^\perp$, and in order to prove (12) it suffices to show that

$$B^+A^+A(A^*H \cap \ker B^*) = \{0\}. \tag{13}$$

Since the orthogonal projection $P = A^+A$ is the identity on $A^*H$ which contains $(A^*H \cap \ker B^*)$, (13) follows from

$$B^+(A^*H \cap \ker B^*) = \{0\},$$

which is clearly a consequence of the definition of $B^+$ and the fact that

$$\ker B^* = (BH)^\perp = \ker B^+.$$

This establishes condition (4).

We assume that the conditions (2), (3) and (4) hold. In order to show that (1) holds, we take $f \in (\ker AB)^\perp$, $g \in (ABH)^\perp$ and we show that

$$B^+A^+(AB)f = f, \tag{14}$$

and

$$B^+A^+g = 0. \tag{15}$$

From (3) it follows that $(\ker A)^\perp \supset B(\ker AB)^\perp$ since $(\ker A)^\perp = A^*H$ and

$$B(\ker AB)^\perp = B(AB)^*H = BB^*A^*H.$$

It follows that $(A^+A)Bf = Bf$ and (14) will be established if we show that $B^+Bf = f$. This last equation is a straightforward consequence of the observation that $(\ker B)^\perp \supset (\ker AB)^\perp$.

In order to show that (15) holds, we note that $AH \supset ABH$ and so $(AH)^\perp \subset (ABH)^\perp$; therefore $(ABH)^\perp$ is the orthogonal direct sum of $(AH)^\perp$ and $AH \cap (ABH)^\perp$. Of course, $B^+A^+$ is identically zero on $(AH)^\perp$ since it is the kernel of $A^+$. Thus it suffices to show that

$$B^+A^+[AH \cap (ABH)^\perp] = \{0\}$$

or

$$A^+[AH \cap (ABH)^\perp] \subset (BH)^\perp. \tag{16}$$

For $f \in A^+[AH \cap \ker(B^*A^*)]$, we know that $Af \in \ker(B^*A^*)$ or $B^*A^*Af = 0$ or $A^*Af \in \ker B^*$. Consequently, $A^*Af \in [A^*H \cap \ker B^*]$. By (4) and the fact that $A^*A$ is self-adjoint, we deduce that $A^*H \cap \ker B^*$ reduces $A^*A$. It is a straightforward consequence of this fact that

$$(A^*A)^{-1}[A^*H \cap \ker B^*] \subset A^*H \cap \ker B^*.$$

We now have established that $f \in A^*H \cap \ker B^*$. It is apparent that (16) holds and consequently we have proved that

$$B^+A^+[(ABH)^\perp] = \{0\}$$

as desired.

Since (2) holds, $ABH$ is closed and $H$ can be written as the orthogonal direct sum of $ABH$ and $(ABH)^\perp$. We have shown that $(AB)^+$ and $B^+A^+$ agree on both of the direct summands. The equation (1) follows.

*Remark* 3.2. The reader will note that the conditions (2), (3), and (4) could not readily be simplified. Condition (3) is equivalent to the agreement of $(AB)^+$ and $B^+A^+$ on $ABH$; condition (4) is equivalent to the agreement of $(AB)^+$ and $B^+A^+$ on $(ABH)^\perp$; condition (2) is a crude necessary condition which is essential in showing that (3) and (4) suffice. A communication with W. S. Loud led the author to discover that Theorem 3.1 remained true when conditions (3) and (4) were replaced with the following two conditions:

$$A^+A \text{ commutes with } BB^*, \tag{3'}$$

$$BB^+ \text{ commutes with } A^*A. \tag{4'}$$

**4. Applications.** Our primary application of Theorem 3.1 is to give a very simple technique for computing the pseudo-inverse of a bounded operator with closed range. The technique is based on a factorization formula for the given operator and the observation that one can easily compute the pseudo-inverse for a normal operator.

LEMMA 4.1. *If $T$ is an operator with closed range, then* $T = (T^*)^+(T^*T)$.

*Proof.* Let $P$ be the orthogonal projection onto $TH = (\ker T^*)^\perp$. We note that

$$T = PT = [(T^*)^+(T^*)]T = (T^*)^+(T^*T).$$

By the closed range theorem of Banach, the fact that $TH$ is closed implies that $T^*H$ is closed. Thus $(T^*)^+$ is bounded and the factorization involves only bounded operators.

THEOREM 4.2. *If $T$ is an operator with closed range, then $T^+ = (T^*T)^+T^*$.*

*Proof.* In order to apply Theorem 3.1 we must show that;

(i) $((T^*)^+)^*H$ is invariant under $(T^*T)^2$,

(ii) $((T^*)^+)^*H \cap \ker(T^*T)$ is invariant under $((T^*)^+)^*(T^*)^+$.

It is clear from the definition of the pseudo-inverse that the kernel of $(T^*)^+$ is $(T^*H)^\perp$, and since the range of $((T^*)^+)^*$ is closed by the closed range theorem, we see that it is $T^*H$. Thus condition (i) above is trivial.

Because $\ker T^* = (TH)^\perp$, we know that $T^*H = T^*(TH) = (T^*T)H$ and since $(T^*T)$ is self-adjoint we have

$$\ker(T^*T) = ((T^*T)H)^\perp = (T^*H)^\perp.$$

By the preceding paragraph, it is clear that the subspace mentioned in (ii) above is trivial. Thus condition (ii) is trivially satisfied, and we may invoke Theorem 3.1 in order to get the desired conclusion.

In order to understand why the preceding theorem gives a simple scheme for computing the pseudo-inverse of any operator $T$ with closed range, the reader should recall the following facts.

PROPOSITION 4.3. *Let $T$ be any operator with closed range. The closed subspace $T^*H$ reduces $(T^*T)$ to an invertible operator $A$, and on the orthogonal complement $(T^*T)$ is the zero operator, which we shall call $B$. Then we have*

$$(T^*T)^+ = A^{-1} \oplus B,$$

*and so we need only compute the inverse of $A$ in order to get the pseudo-inverse of $(T^*T)$.*

In the most frequent applications of the pseudo-inverse, the operator $T$ is defined on a finite-dimensional vector space. In that situation, it is entirely straightforward to calculate $A^{-1}$ given $A$. Thus in order to give a complete technique for computing the pseudo-inverse in that situation, we need only describe how to obtain a matrix for $A$. Such a problem is clearly equivalent to finding an orthonormal basis for $T^*H$. If one takes any convenient basis for $H$ and writes the corresponding matrix for $T^*$ so that the matrix acts on a column vector written to the right of the matrix, then the row vectors of the matrix span $T^*H$. If one applies the Gram–Schmidt orthonormalization process to those row vectors, and if one denotes by $\{e_1, \cdots, e_j\}$ the orthonormal set obtained from $\{a_1, \cdots, a_i\}$, then it is easy to determine if $a_{i+1}$ is linearly dependent on the set $\{e_1, \cdots, e_j\}$. If the linear combination $c_1e_1 + \cdots + c_je_j$ with $c_k = \langle a_{i+1}, e_k\rangle$ equals $a_{i+1}$, then it is linearly dependent on the set, and otherwise it is not. Thus the Gram–Schmidt process results in a maximal orthonormal set for $T^*H$, and obtaining the matrix for $(T^*T)$ in that basis is routine. Hence, we have given a complete straightforward procedure for computing the pseudo-inverse, and our procedure seems simpler and easier than the previously known procedures.

A useful tool in operator theory is the polar factorization (see [7, pp. 1245–1250]). Consequently the following theorem, which is almost immediate from Theorem 3.1, may prove useful in general operator theory.

THEOREM 4.4. *Let $UR$ be the polar factorization of the operator $T$. If $R$ has closed range, then $T^+$ is the composite operator $R^+U^*$.*

*Proof.* By the usual construction of the polar factorization, $U$ is a partial isometry carrying $RH$ isometrically onto $TH$. Thus if $RH$ is closed, then $TH$ is closed and so $R^+$ and $T^+$ are both bounded. Any partial isometry has closed range and so $U^+$ is bounded; moreover, it is well known that $U^+ = U^*$.

One quickly sees that $U^*H$ is $RH$ which is obviously invariant under the self-adjoint operator $R$. Since the kernel of $R$ is $(RH)^\perp$, the subspace $U^*H \cap \ker R$ is certainly trivial. It is now clear that the conditions (2), (3), and (4) of our Theorem 3.1 are satisfied.

REFERENCES

[1] F. J. BEUTLER, *The operator theory of the pseudo-inverse, I*, J. Math. Anal. Appl., 10 (1965), pp. 451–470.
[2] ———, *The operator theory of the pseudo-inverse, II*, Ibid., 10 (1965), pp. 471–493.
[3] R. H. BOULDIN, *The product of two operators with losed range*, Tôhoku Math. J., to appear.
[4] T. L. BOULLION AND P. L. ODELL, *Generalized Inverse Matrices*, Wiley-Interscience, New York, 1971.
[5] C. DESOER AND B. WHALEN, *A note on pseudo-inverses*, J. Soc. Indust. Appl. Math., 11 (1963), pp. 442–447.
[6] R. G. DOUGLAS, *On majorization, factorization, and range inclusion of operators on Hilbert space*, Amer. Math. Soc., 17 (1966), pp. 413–415.
[7] N. DUNFORD AND J. T. SCHWARTZ, *Linear Operators*, vol. II, Interscience, New York, 1963.
[8] T. N. E. GREVILLE, *The pseudo-inverse of a rectangular or singular matrix and its application to the solution of systems of linear equations*, J. Soc. Indust. Appl. Math., 1 (1959), pp. 38–43.
[9] ———, *Some applications of the pseudo-inverse of a matrix*, Ibid., 2 (1960), pp. 15–22.
[10] P. R. HALMOS, *Introduction to Hilbert Space*, Chelsea, New York, 1957.
[11] T. KATO, *Perturbation Theory for Linear Operators*, Springer–Verlag, New York, 1966.