I. Introduction. The theory of pseudoinverses, or generalized inverses, has been extensively developed over the last few years. A recent and comprehensive bibliography occurs in [2] which gives a short history also, and a brief survey of known results and computational methods.

We consider here the case of a bounded operator with closed range between Hilbert spaces. Most of the information for this case, which is used in this paper can be found in [3], which is pithy and short.

II. Notation and basic information. \( A \) is a bounded linear transformation with closed range from \( S_1 \) to \( S_2 \), where \( S_1 \) and \( S_2 \) are Hilbert spaces with inner product norms \( | \cdot |_1 \) and \( | \cdot |_2 \), respectively. \( A^* \) is the adjoint of \( A \), \( A^\dagger \) is the pseudoinverse of \( A \) (uniquely defined for this case [3]). \( I_1 \) and \( I_2 \) are the identity mappings on \( S_1 \) and \( S_2 \), respectively. \( \overline{S}_1 = A^*S_2 = A^*A S_1 \) and \( \overline{S}_2 = AS_1 = AA^*S_2 \). We use a bar to denote the restriction of a mapping to \( \overline{S}_1 \) or \( \overline{S}_2 \), as appropriate, e.g., \( \overline{A} = A|_{\overline{S}_1} \) the restriction of \( A \) to \( \overline{S}_1 \). We know that \((\overline{A})^{-1}\) exists and is bounded [3]; and it is readily deduced that \( |A^\dagger| = |(\overline{A})^{-1}| \) (from [3, Theorem 3], noting that \( \overline{S}_1 = \mathfrak{N}(A)^\perp_1 \), and \( \overline{A}^\dagger - (\overline{A})^{-1} = 0 \)). Thus, we know a priori that \( A^\dagger \) is a uniquely defined, bounded linear transformation from \( S_2 \) to \( S_1 \). We use the conventional norms on linear transformations. We use the basic definition of the pseudoinverse given by Desoer and Whalen [3], for the case studied \( B = A^\dagger \) if and only if \( B \) is a linear operator such that

(i) \( B\overline{A} = \overline{I}_1 \),
(ii) \( B\overline{S}_2^\perp = 0 \) (i.e., \( By = 0 \) if \( y \) is in the orthogonal complement of \( \overline{S}_2 \)).

From this definition it follows that \( AA^\dagger A = A \) and \( A^*AA^\dagger = A^* \), since \( S_1 = \overline{S}_1 \oplus \overline{S}_1^\perp \) and \( S_2 = \overline{S}_2 \oplus \overline{S}_2^\perp \).

III. Theorem 1. Define, for \( t \geq 0 \),

\[
A^\dagger(t) = \int_0^t \exp[-A^*A(t-s)]A^*ds.
\]
Then

\[ |A^t - A^t(t)| \leq |A^t| \exp[-t|A^t|^{-2}], \quad \forall t. \]

**Proof.** Since \( A^* = A^*AA^t \), \( (3.1) \) may be evaluated as: \( A^t(t) = [I_1 - \exp\{-tA^*A\}]A^t \). Pick \( y \in S_2 \), and let \( \alpha(t) = [A^t - A^t(t)]y = \exp\{-tA^*A\}A^t y \). Clearly, \( \alpha(t) \in S_1, \forall t \). Thus \( d|\alpha(t)|^2/\!\!dt = -2|A\alpha(t)|^2 \) (see [5]) \( \leq -2|A^t|^{-2} |\alpha(t)|^2 \), whence \( |A^t y - A^t(t)y| \leq |A^t y| \exp[-t|A^t|^{-2}] \), and \( (3.2) \) follows since \( y \) is arbitrary.

The next theorem has been proved for the matrix case by den Broeder and Charnes (see [2]), and for the nonsingular case by Altman [1].

**Theorem 2.** Pick \( c \) so that \( 0 < c < 2, \forall n \geq 0 \), define

\[ B_{n,c} = \sum_{p=0}^n \left[ I_1 - \frac{cA^*A}{|A|^2} \right]^p \frac{cA^*}{|A|^2}. \]

Define:

\[ \beta_e = \max\{ |1 - c|, |1 - c/|A|^2|, |A^t|/|A|^2 | \} < 1. \]

Then

\[ |A^t - B_{n,c}| \leq (c\beta_e^{n+1}/(1 - \beta_e)) |A|, \quad \forall n \geq 0. \]

**Proof of Theorem 2.** Let \( B_{\infty,c} = \lim_{n \to \infty} B_{n,c} \), and define \( E = cA^*A/|A|^2 \). Then \( |I_1 - E| \leq \beta_e \) (see [1, pp. 52–55]) so that

\[ (E)^{-1} = \sum_{p=0}^\infty (I_1 - E)^p \text{ (where } (E)^{-1} \text{ is defined on } S_1, \text{ the range of } E). \]

Now \( E: S_1 \to S_1 \) and \( (I_1 - E): S_1 \to S_1 \), so

\[ B_{\infty,c} = \sum_{p=0}^\infty (I_1 - E)^p \overline{E} = \sum_{p=0}^\infty (I_1 - E)^p \overline{E} = \overline{I_1}. \]

\( B_{n,c} = 0, \forall n, \) since \( A^*S_2 = 0, \) in view of (3.3). Linearity is obvious. Thus \( A^t = B_{\infty,c} \); and (3.4) follows immediately, since it is the rate of convergence on \( S_2 \) (see [1, p. 52–55]). Optimal convergence occurs when

\[ c = 2 |A|^2 |A^t|^2/(|A|^2 |A^t|^2 + 1), \]

in which case

\[ \beta_e = (|A|^2 |A^t|^2 - 1)/(|A|^2 |A^t|^2 + 1) = [1], \]

which yields the following
Corollary to Theorem 2. If \( |A| = |A^\dagger| = 1 \), then \( A^\dagger = A^* \).

The rates of convergence in the previous theorems are often impracticably slow. The following theorem gives a rapidly convergent recursion sequence, analogous to a well-known technique for improving estimates of the inverse for a nonsingular matrix [4, p. 120].

**Theorem 3.** Pick \( c \) and define \( \beta_c \) as in Theorem 2. Define the sequence:

\[
D_{0,c} = cA^*/|A|^2; \quad D_{n+1,c} = 2D_{n,c} - D_{n,c}AD_{n,c}.
\]

Then

\[
|A^\dagger - D_{n,c}| \leq |A^\dagger| \cdot \beta_c^{2^n}.
\]

**Proof of Theorem 3.** Define \( \Gamma_c = [A^\dagger A - D_{0,c}A] \). \( \Gamma_cS_1 = 0 \) and \( \Gamma_c = I_1 - E \) (see proof of Theorem 2) so \( |\Gamma_c| \leq \beta_c \). By definition \( \Gamma_c^{2^n} = [A^\dagger A - D_{0,c}A] \). Suppose that \( \Gamma_c^{2^n} = [A^\dagger A - D_{n,c}A] \). Then

\[
\Gamma_c^{2^n+1} = (\Gamma_c^{2^n})^2 = A^\dagger AA^\dagger A - A^\dagger AD_{n,c}A
\]

\[
- D_{n,c}AA^\dagger A + D_{n,c}AD_{n,c}A
\]

\[
= A^\dagger A - (2D_{n,c} - D_{n,c}AD_{n,c})A
\]

\[
= [A^\dagger A - D_{n+1,c}A],
\]

since \( D_{n,c}A : S_1 \rightarrow S_1 \), \( \forall n \), which is clear from (3.5), and since \( AA^\dagger A = A \). Also from (3.5), \( D_{n,c}S_2 = 0 = A^\dagger S_2 \), \( \forall n \). Since \( |\Gamma_c| \leq \beta_c \), and \( A^\dagger - D_{n,c} = (A^\dagger - D_{n,c})AA^\dagger \), \( \forall n \), (3.6) follows.

IV. So long as \( |A^\dagger| < \infty \), the hypothesis that \( |A| < \infty \) may be relaxed in Theorem 1. We need only require that \( A^*A \) is a closed mapping with dense domain. The proof as given remains valid (see [5]).

If \( A \) fails to have closed range, but is bounded, then \( A^\dagger(t) \), \( B_{n,c} \), and \( D_{n,c} \) converge monotonically to \( A^\dagger \), but (3.2), (3.4), (3.6) are useless since \( |A^\dagger| = \infty \) and \( \beta_c = 1 \). The associated proofs are more lengthy and will be published elsewhere.

**References**


**Courant Institute of Mathematical Sciences, New York University**