New perturbation results on pseudo-inverses of linear operators in Banach spaces

Jiu Ding *

Department of Mathematics, The University of Southern Mississippi, Hattiesburg, MS 39406-5045, USA
Received 24 October 2001; accepted 5 July 2002
Submitted by R.A. Brualdi

Abstract

Let $X$ and $Y$ be Banach spaces, let $T : X \to Y$ be a bounded linear operator with closed range, and let $S : X \to Y$ be another bounded linear operator. We study conditions on $S - T$ that guarantee the closeness of the range of $S$ and obtain some new bounds on the pseudo-inverse of $S$.

© 2002 Published by Elsevier Science Inc.

AMS classification: 15A09; 65F20

Keywords: Pseudo-inverses; Perturbations

1. Introduction

Perturbation analysis for pseudo-inverses of bounded linear operators of Banach spaces is very important in practical applications of operator theory and has been widely studied, see [11]. In recent years the perturbation study of pseudo-inverses with the help of the concept of the gap between closed subspaces [10] has appeared in [3,4,8,9]. Especially in the work in [2,5], such perturbation results have been applied to frame theory.

Motivated by the ideas in [2,5], we further explore the following general question: Let $T$ and $S$ be two bounded linear operators from a Banach space $X$ to a Banach space $Y$ such that the range of $T$ is closed, so that the pseudo-inverse $T^\dagger$ is well-defined.
What conditions on the difference of $S$ and $T$ guarantee that the range of $S$ is also closed so that the pseudo-inverse $S^\dagger$ is also well-defined, and if so, what is an upper bound of the norm of $S^\dagger$ in terms of that of $T^\dagger$?

A classic result is the Neumann lemma which says that if $P$ is a bounded linear operator on $X$ such that $\|P\| < 1$, then $I + P$ is bijective, and

$$\|(I + P)^{-1}\| \leq \frac{1}{1 - \|P\|}. \quad (1)$$

The assumption of this celebrated result was weakened in [2] in which the authors proved that if there are $\lambda_1, \lambda_2 \in [0, 1)$ such that

$$\|Px\| \leq \lambda_1\|x\| + \lambda_2\|(I + P)x\| \quad \forall x \in X,$$

$I + P$ is bijective, and

$$\|(I + P)^{-1}\| \leq \frac{1 + \lambda_2}{1 - \lambda_1}. \quad (2)$$

In this paper we extend the results on invertible operators to general ones. In particular we give the perturbation results on pseudo-inverses. In the next section we review some concepts and give a basic lemma. In Section 3 we prove the main results. We conclude with Section 4.

2. Preliminaries

Let $X$ and $Y$ be two Banach spaces. Let $B(X, Y)$ be the Banach space of all bounded linear operators $T : X \to Y$ with norm $\|T\| = \sup\{\|Tx\| : \|x\| = 1\}$, and let $B_c(X, Y)$ be the subspace of all $T \in B(X, Y)$ such that the range $R(T)$ of $T$ is closed in $Y$. If $X = Y$, we write $B(X, X)$ and $B_c(X, X)$ as $B(X)$ and $B_c(X)$, respectively.

Let $T \in B_c(X, Y)$. Throughout the paper we assume that $X$ is the topological sum of the null space $N(T)$ of $T$ and $N(T)^c$, and $Y$ is the topological sum of $R(T)$ and $R(T)^c$, where $N(T)^c$ and $R(T)^c$ are closed subspaces of $X$ and $Y$, respectively. Note that $T$ is one-to-one from $N(T)^c$ onto $R(T)$. Let $P$ be the projection of $X$ onto $N(T)$ along $N(T)^c$, and let $Q$ be the projection of $Y$ onto $R(T)$ along $R(T)^c$. The bounded linear operator $T^\dagger : Y \to X$ defined by $T^\dagger Ty = x$ for $x \in N(T)^c$ and $T^\dagger y = 0$ for $y \in R(T)^c$ is called the pseudo-inverse of $T$ (with respect to $P, Q$) [1,11]. If $X$ and $Y$ are Hilbert spaces, and if $N(T)^c = N(T)^\perp$ and $R(T)^c = R(T)^\perp$, then the corresponding pseudo-inverse of $T$ is usually referred to as the Moore–Penrose pseudo-inverse.

We also need the concept of the approximate point spectrum of a bounded linear operator $T$ to prove Lemma 2.2. A complex number $\alpha$ is said to be in the approximate point spectrum $\sigma_a(T)$ of $T$ if there exists a sequence $x_n$ of vectors such that $\|x_n\| = 1$ for all $n$ and $\|(\alpha I - T)x_n\| \to 0$. It is obvious that if there is a positive number $\varepsilon$ such
that \( \| (\alpha I - T)x \| \geq \epsilon \| x \| \) for all vectors \( x \), then \( \alpha \notin \sigma_a(T) \). The following lemma is a standard result concerning \( \sigma_a(T) \) [6, Proposition VII.6.7].

**Lemma 2.1.** \( \sigma_a(T) \supset \partial \sigma(T) \), the boundary of the spectrum \( \sigma(T) \) of \( T \).

The following lemma is itself a key perturbation result which directly generalizes the Neumann lemma (1) and will be used in proving Corollary 3.2 and Proposition 3.1.

**Lemma 2.2.** Let \( P \in B(X) \) be such that
\[
\| Px \| \leq \lambda_1 \| x \| + \lambda_2 \| (I + P)x \| \quad \forall x \in X,
\]
where \( \lambda_1 < 1 \) and \( \lambda_2 < 1 \). Then \( \lambda_1 \in (-1, 1) \), \( \lambda_2 \in (-1, 1) \), and \( I + P \) is bijective. Moreover,
\[
\frac{1 - \lambda_1}{1 + \lambda_2} \| x \| \leq \| (I + P)x \| \leq \frac{1 + \lambda_1}{1 - \lambda_2} \| x \| \quad \forall x \in X,
\]
\[
\frac{1 - \lambda_2}{1 + \lambda_1} \| y \| \leq \| (I + P)^{-1}y \| \leq \frac{1 + \lambda_2}{1 - \lambda_1} \| y \| \quad \forall y \in X.
\]

**Proof.** From
\[
\| (I + P)x \| \geq \| x \| - \| Px \| \geq \| x \| - \lambda_1 \| x \| - \lambda_2 \| (I + P)x \| \quad \forall x \in X,
\]
\[
(1 + \lambda_2)\| (I + P)x \| \geq (1 - \lambda_1)\| x \|
\]
from which \( 1 + \lambda_2 > 0 \) and the left inequality of (4) follows. Since
\[
\| (I + P)x \| \leq \| x \| + \| Px \| \leq (1 + \lambda_1)\| x \| + \lambda_2\| (I + P)x \| \quad \forall x \in X.
\]
\[
(1 - \lambda_2)\| (I + P)x \| \leq (1 + \lambda_1)\| x \|
\]
Hence \( 1 + \lambda_1 > 0 \) and the right inequality of (4) is obtained.

In order to prove (5), thanks to (4), it is enough to show that \( -1 \notin \sigma(P) \). Let \( \alpha \leq -1 \). Then the triangle inequality and (3) imply that
\[
\| \alpha x - Px \| \geq -\| x \| - \| Px \| \geq -(\alpha + 1)\| |x|\| - \lambda_2\| (I + P)x \|.
\]

We consider \( 0 \leq \lambda_2 < 1 \) and \( -1 < \lambda_2 < 0 \) separately. If \( 0 \leq \lambda_2 < 1 \), then
\[
\| \alpha x - Px \| \geq -(\alpha + \lambda_1)\| x \| - \lambda_2\| (\| \alpha x - Px \| - (\alpha + 1)\| x \|) = -\alpha - \lambda_1 + \lambda_2(\alpha + 1)\| x \| - \lambda_2\| \alpha x - Px \|.
\]
Hence for \( \epsilon = \left[ -\alpha - \lambda_1 + \lambda_2(\alpha + 1) \right]/(1 + \lambda_2) > 0 \),
\[
\| \alpha x - Px \| \geq \epsilon \| x \| \quad \forall x \in X.
\]
If \( -1 < \lambda_2 < 0 \), then for \( \epsilon = -(\alpha + \lambda_1) > 0 \),
\[
\| \alpha x - Px \| \geq \epsilon \| x \| \quad \forall x \in X.
\]
Thus we see that \( \alpha / \sigma (P) \) for any \( \alpha \leq -1 \). Since \( \alpha \notin \sigma (P) \) for all \( \alpha < -\|P\| \), if \(-1 \in \sigma (P)\), there must be some \( \alpha_0 \leq -1 \) such that \( \alpha_0 \in \sigma (P) \). Therefore \( \alpha_0 \in \sigma (P) \) by Lemma 2.1, which leads to a contradiction. 

Remark 2.1. Lemma 2.2 under a stronger condition that \( \lambda_1, \lambda_2 \in [0, 1) \) was also proved in [2] with the help of the concept of dual operators. Our direct proof here is much simpler.

3. Perturbation of bounded linear operators

Now we study the perturbation problem for pseudo-inverses of general bounded linear operators of Banach spaces. In the following we exclude the trivial case of \( T = 0 \).

Theorem 3.1. Let \( T, S \in B(X, Y) \) be such that

\[
\|(S - T)x\| \leq \lambda_1 \|Tx\| + \lambda_2 \|Sx\| \quad \forall x \in X,
\]

where \( \lambda_1 < 1 \). Then \( \lambda_2 > -1 \) and

\[
\|Sx\| \geq \frac{1 - \lambda_1}{1 + \lambda_2} \|Tx\| \quad \forall x \in X.
\]

If in addition \( N(S) = N(T) \), then \( T \in B_c(X, Y) \) implies \( S \in B_c(X, Y) \), and in this case,

\[
\|S^\dagger\| \leq \frac{1 + \lambda_2}{1 - \lambda_1} \|T^\dagger\|.
\]

Proof. From

\[
\|Sx\| \geq \|Tx\| - \|Sx - Tx\| \geq \|Tx\| - \lambda_1 \|Tx\| - \lambda_2 \|Sx\|,
\]

\((1 + \lambda_2)\|Sx\| \geq (1 - \lambda_1)\|Tx\|\), and so \( \lambda_2 > -1 \) and (7) follows.

Now assume further that \( N(S) = N(T) \). Let \( y_n = Sx_n \in R(S) \) converge to \( y \in Y \), where without loss of generality, \( x_n \in N(S)^c = N(T)^c \). Let \( z_n = Tx_n \). Then from

\[
\|y_n\| \geq \frac{1 - \lambda_1}{1 + \lambda_2} \|z_n\|,
\]

we see that \( z_n \) is a Cauchy sequence, and so \( z_n \) converges to some vector \( z \in R(T) \).

Since \( x_n = T^\dagger z_n \) and since \( T^\dagger \in B(Y, X) \), \( x_n \) converges to some \( x \in X \). Now \( Sx_n = y_n \) converges to \( Sx = y \in R(S) \). This proves that \( R(S) \) is closed.

To prove the inequality (8), let \( y \in R(S) \). Then \( y = Sx \) for some \( x \in N(S)^c \), and so \( x = S^\dagger y \). Since \( N(S)^c = N(T)^c \), we have \( T^\dagger T S^\dagger y = S^\dagger y \). Therefore, (7) implies that

\[
\|y\| \geq \frac{1 - \lambda_1}{1 + \lambda_2} \|T S^\dagger y\| \geq \frac{1 - \lambda_1}{1 + \lambda_2} \|T^\dagger T S^\dagger y\| = \frac{1 - \lambda_1}{1 + \lambda_2} \|S^\dagger y\|.
\]
which gives that
\[ \|S^\dagger y\| \leq \frac{1 + \lambda_2}{1 - \lambda_1} \|T^\dagger\| \|y\| \quad \forall y \in Y. \]

\[ \square \]

**Remark 3.1.** In the context of Hilbert spaces, Theorem 3.1 was also obtained in [5, Theorem 2.2] using the concept of the reduced minimum modulus of linear operators.

Let \( T = I \) and \( S = I + P \). Then Theorem 3.1 gives:

**Corollary 3.1.** Let \( P \in B(X) \) be such that (3) is valid with \( \lambda_1 < 1 \). Then \( \lambda_2 > -1 \), \( I + P \in B_c(X) \) and is injective. Moreover,

\[ \|(I + P)x\| \geq \frac{1 - \lambda_1}{1 + \lambda_2} \|x\| \quad \forall x \in X, \quad (9) \]

\[ \|(I + P)^\dagger y\| \leq \frac{1 + \lambda_2}{1 - \lambda_1} \|y\| \quad \forall y \in X. \quad (10) \]

**Remark 3.2.** Corollary 3.1 extends (2) from invertible cases to general cases.

**Theorem 3.2.** Let \( T \in B_c(X, Y) \), \( S \in B(X, Y) \) be such that (6) is satisfied with \( \lambda_1 < 1 \) and \( \lambda_2 < 1 \). Then \( \lambda_1 \in (-1, 1) \) and \( \lambda_2 \in (-1, 1) \). Furthermore, \( S \in B_c(X, Y) \), \( N(S) = N(T) \), and

\[ \frac{1 - \lambda_1}{1 + \lambda_2} \|Tx\| \leq \|Sx\| \leq \frac{1 + \lambda_1}{1 - \lambda_2} \|Tx\| \quad \forall x \in X, \quad (11) \]

\[ \|S^\dagger y\| \geq \frac{1 - \lambda_2}{1 + \lambda_1} \|T\| \|y\| \quad \forall y \in R(S), \quad (12) \]

\[ \|S^\dagger y\| \leq \frac{1 + \lambda_2}{1 - \lambda_1} \|T^\dagger\| \|y\| \quad \forall y \in Y. \quad (13) \]

**Proof.** The assertion that \( \lambda_1 \in (-1, 1) \) and \( \lambda_2 \in (-1, 1) \) comes from the same argument as in the proof of Lemma 2.2. The left inequality of (11) and the inequality (13) have already been proved in Theorem 3.1. The right inequality of (11) follows from the fact that

\[ \|Sx\| \leq \|Sx - Tx\| + \|Tx\| \leq (1 + \lambda_1)\|Tx\| + \lambda_2\|Sx\|. \]

Eq. (11) implies that \( N(S) = N(T) \), and so by Theorem 3.1, \( R(S) \) is closed. To prove (12), let \( Sx = y \) with \( x \in N(S)^c \) and \( y \in R(S) \). Then \( x = S^\dagger y \) and the right inequality of (11) gives that

\[ \|y\| \leq \frac{1 + \lambda_1}{1 - \lambda_2} \|T S^\dagger y\| \leq \frac{1 + \lambda_1}{1 - \lambda_2} \|T\| \|S^\dagger y\|, \]

and so (12) follows. \[ \square \]
Corollary 3.2. Under the same assumptions as in Theorem 3.2, if in addition $T$ is bijective, then so is $S$. In this case,
\[
\frac{1 - \lambda_2}{1 + \lambda_1 \|T\|} \|y\| \leq \|S^{-1}y\| \leq \frac{1 + \lambda_2}{1 - \lambda_1} \|T^{-1}\| \|y\| \quad \forall y \in Y.
\] (14)

**Proof.** Because of Theorem 3.2, it is enough to show that $S$ is bijective. Let $L = ST^{-1}$ and $x = T^{-1}y$. Then (6) implies that
\[
\|Ly - y\| \leq \lambda_1 \|y\| + \lambda_2 \|Ly\| \quad \forall y \in Y.
\]
By Lemma 2.2, $L$ is bijective, and so $S = LT$ is bijective. $\square$

Remark 3.3. Corollary 3.2 was also proved in [2] under a stronger condition that $\lambda_1, \lambda_2 \in [0, 1)$.

Remark 3.4. Lemma 2.2 is a special case of Corollary 3.2 with $T = I$ and $S = I + P$.

Now we give an expression for $S^{\dagger}$ if $T$ is surjective.

**Proposition 3.1.** Under the same assumptions as in Theorem 3.2, if in addition $T$ is surjective, then so is $S$, and
\[
S^{\dagger} = T^{\dagger}(I + AT^{\dagger})^{-1},
\] (15)
where $A = S - T$. In particular, if $T$ is bijective, then
\[
S^{-1} = T^{-1}(I + AT^{-1})^{-1}.
\] (16)

**Proof.** Since $N(S) = N(T)$, we have
\[
S = T + A = (I + AT^{\dagger})T.
\]
Since $R(T) = Y$, we see that for all $y \in Y$,
\[
\|AT^{\dagger}y\| = \|(S - T)T^{\dagger}y\| \leq \lambda_1 \|T T^{\dagger}y\| + \lambda_2 \|ST^{\dagger}y\|
\leq \lambda_1 \|y\| + \lambda_2 \|(I + AT^{\dagger})y - (I - TT^{\dagger})y\|
\leq \lambda_1 \|y\| + \lambda_2 \|(I + AT^{\dagger})y\|.
\]
Hence, $I + AT^{\dagger}$ is bijective by Lemma 2.2, which implies that $R(S) = R(T) = Y$. Finally the expression (15) follows from Lemma 3.2 of [8] since $T$ is surjective. $\square$

4. Conclusions

In this paper we extended some of the previous perturbation results for invertible linear operators to arbitrary ones and obtained some new upper bounds for pseudo-inverses of perturbed operators of some kind, which weakens the assumption on the
norm of the perturbation. Such results have direct applications to error estimates to least squares solutions of linear operator equations with weaker assumptions, after they are combined with some well-known (see, e.g., [7]) perturbation theorems.

Acknowledgement

We thank the referee for good suggestions on the re-organization of the paper.

References