Frame Approximation of Pseudo-Inverse Operators

Ole Christensen

Department of Mathematics, Technical University of Denmark,
Building 303, 2800 Lyngby, Denmark
E-mail: Ole.Christensen@mat.dtu.dk

Submitted by J. H. Shapiro

Received September 14, 1999

Let $T$ denote an operator on a Hilbert space $(\mathcal{H}, \langle \cdot, \cdot \rangle)$, and let \{\phi_n\}_{i=1}^\infty be a frame for the orthogonal complement of the kernel $N_T$. We construct a sequence of operators \{\Phi_n\}_{i=1}^\infty which converges to the pseudo-inverse $T^\dagger$ of $T$ in the strong operator topology as $n \to \infty$. The operators \{\Phi_n\}_{i=1}^\infty can be found using finite-dimensional methods. We also prove an adaptive iterative version of the result. © 2001 Academic Press

1. LINEAR APPROXIMATION OF $T^\dagger f$

Let $\mathcal{H}$ be a separable Hilbert space. Let $T : \mathcal{H} \to \mathcal{H}$ be a bounded linear operator with closed range $R_T$. Denote the kernel of $T$ by $N_T$.

Let us recall the definition of the pseudo-inverse of $T$. By restricting $T$ to the orthogonal complement of the kernel, $N_T^\perp$, we obtain a bijective operator

$$\tilde{T} : N_T^\perp \to R_T.$$

$\tilde{T}$ has a bounded inverse $\tilde{T}^{-1} : R_T \to N_T^\perp$. The pseudo-inverse of $T$ is now defined as the unique extension $T^\dagger$ of $\tilde{T}^{-1}$ to an operator from $\mathcal{H}$ into $\mathcal{H}$ for which $N_{T^\dagger} = R_{T^\dagger}$.

Alternatively, $T^\dagger$ can be characterized as the unique linear operator from $\mathcal{H}$ to $\mathcal{H}$ for which

$$N_{T^\dagger} = R_{T^\dagger}, \quad (1)$$
$$R_{T^\dagger} = N_{T^\dagger}, \quad (2)$$
$$TT^\dagger f = f, \quad \forall f \in R_T. \quad (3)$$
It is well known that $TT^\dagger$ is the orthogonal projection of $\mathcal{H}$ onto $R_T$ and that $T^\dagger T$ is the orthogonal projection of $\mathcal{H}$ onto $N^\perp_T$. We refer to [2] for those facts and other results concerning pseudo-inverses. Let $P_{R_T}$ denote the orthogonal projection of $\mathcal{H}$ onto $R_T$ and observe that for arbitrary $f \in \mathcal{H}$ we have $(I - P_{R_T})f \in R^\perp_T = N^\perp_T$; therefore

$$T^\dagger f = T^\dagger P_{R_T} f + T^\dagger (I - P_{R_T}) f = T^\dagger P_{R_T} f.$$ 

The purpose of this note is to present a method for approximation of $T^\dagger$. Since the method is based on a frame decomposition of $N^\perp_T$ we first recall the basic facts from frame theory.

A family of elements $\{f_i\}_{i \in I} \subseteq \mathcal{H}$ is a frame for $\mathcal{H}$ if there exist constants $A,B > 0$ such that

$$A\|f\|^2 \leq \sum_{i \in I} |\langle f, f_i \rangle|^2 \leq B\|f\|^2, \quad \forall f \in \mathcal{H}. \quad (4)$$

The numbers $A,B$ are called frame bounds. If $\{f_i\}_{i \in I}$ is a frame, the frame operator is

$$S: \mathcal{H} \to \mathcal{H}, \quad Sf = \sum_{i \in I} \langle f, f_i \rangle f_i. \quad (5)$$

The fact that $S$ is a bounded, invertible, and positive operator on $\mathcal{H}$ leads to the frame decomposition,

$$f = SS^{-1} f = \sum_{i \in I} \langle f, S^{-1} f_i \rangle f_i, \quad \forall f \in \mathcal{H}. \quad (6)$$

So every element $f \in \mathcal{H}$ has a representation as an (infinite) linear combination of the frame elements. In particular, a frame for $\mathcal{H}$ is total in $\mathcal{H}$, i.e., $\text{span}\{f_i\}_{i \in I} = \mathcal{H}$. However, if $\{f_i\}_{i \in I}$ is not total, $\{f_i\}_{i \in I}$ can still be a frame for the subspace $\text{span}\{f_i\}_{i \in I}$, in which case we call $\{f_i\}_{i \in I}$ a frame sequence. In particular, every finite set of vectors $\{f_i\}_{i \in I}$ is a frame sequence.

A frame for a subspace yields a representation of the orthogonal projection onto the subspace. Given a sequence $\{f_i\}_{i \in I}$, let $P_I$ denote the orthogonal projection onto $\text{span}\{f_i\}_{i \in I}$. The lemma below appeared in [4].

**Lemma 1.1.** Let $\{f_i\}_{i \in I}$ be a frame sequence. Then the orthogonal projection onto $\text{span}\{f_i\}_{i \in I}$ is given by

$$P_I f = \sum_{i \in I} \langle f, S^{-1} f_i \rangle f_i, \quad f \in \mathcal{H}.$$ 

Throughout the paper we will need two “types” of orthogonal projections, namely,

(i) Projections onto a subset of the range of the operator $T$. A projection of this type will be denoted by $P_W$, where $W$ is the subspace we project onto.
(ii) Projections onto the closed span of a subset \( \{f_i\}_{i \in J} \) of the frame \( \{f_i\}_{i \in I} \). This projection will be denoted by \( P_J \).

Now, corresponding to the given operator \( T \), let \( \{f_i\}_{i \in I} \) be a frame for the subspace \( N_T^⊥ \). For each finite subset \( J \subseteq I \), let \( H_J := \text{span}\{f_i\}_{i \in J} \). Then \( \{f_i\}_{i \in J} \) is a frame for \( H_J \) with frame operator given by

\[
S_J : H_J \to H_J, \quad S_J f = \sum_{i \in J} \langle f, f_i \rangle f_i.
\]

Also, \( \{Tf_i\}_{i \in J} \) is a frame for \( TH_J \), with frame operator \( V_J : TH_J \to TH_J \) given by

\[
V_J f = \sum_{i \in J} \langle f, Tf_i \rangle Tf_i
= T \sum_{i \in J} \langle T^* f, f_i \rangle f_i
= TS_J P_J T^* f, \quad f \in TH_J.
\]

i.e.,

\[
V_J = TS_J P_J T^*.
\]

When in the following we write \( V_J^{-1} \), it is understood that we invert \( V_J \) as an operator from \( TH_J \) onto \( TH_J \).

**Lemma 1.2.** Fix \( f \in H \) and let \( J \subseteq I \) be finite. Then \( \inf_{\phi \in H_J} \| f - T\phi \| \) is attained for

\[
\phi = \sum_{i \in J} \langle f, V_J^{-1} f_i \rangle f_i,
\]

and for this choice of \( \phi \),

\[
\| f \|^2 = \| T\phi \|^2 + \| (I - P_{TH_J}) f \|^2.
\]

**Proof.** The element \( \psi \in TH_J \) closest to \( f \), by Lemma 1.1,

\[
\psi = P_{TH_J} f = \sum_{i \in J} \langle f, V_J^{-1} f_i \rangle f_i.
\]

Clearly \( \phi := \sum_{i \in J} \langle f, V_J^{-1} f_i \rangle f_i \) satisfies \( T\phi = \psi \). Further, \( f = P_{TH_J} f + (I - P_{TH_J}) f \) and \( P_{TH_J} f = T\phi \), so the rest of the lemma follows. \( \square \)

Lemma 1.2 leads to a method for approximation of \( T^† f \). Let \( \{I_n\}_{n=1}^\infty \) be a family of finite subsets of \( I \) such that \( I_1 \subseteq I_2 \cdots \subseteq I_n \uparrow I \). Abusing the notation, we will write \( H_n, S_n, V_n, P_n \) instead of \( H_{I_n}, S_{I_n}, V_{I_n}, P_{I_n} \). The following lemma states that for \( f \in R_T \) we can make \( \inf_{\phi \in H_n} \| f - T\phi \| \) arbitrarily small by choosing \( n \) large enough.
Lemma 1.3. Let \( f \in R_T \). Given \( \varepsilon > 0 \), there exists \( N \) such that for all \( n \geq N \),
\[
\inf_{\phi \in \mathcal{H}_n} \| f - T\phi \| \leq \varepsilon.
\]

Proof. Observe that given \( f \in R_T \), \( TP_n T^\dagger f \to f \) for \( n \to \infty \). The result then follows since \( P_n T^\dagger f \in \mathcal{H}_n \).

The next theorem shows how we can obtain a family of operators \( \{ \Phi_n \} \) that converges to \( T^\dagger \) in the strong operator topology.

**Theorem 1.4.** For \( n \in \mathbb{N} \), define \( \Phi_n: \mathcal{H} \to \mathcal{H} \) by
\[
\Phi_n f := \sum_{i \in I_n} \langle f, V_n^{-1} T f_i \rangle f_i.
\]
Then \( \Phi_n f \to T^\dagger f \) for \( n \to \infty \), for all \( f \in \mathcal{H} \).

Proof. Let \( f \in \mathcal{H} \). From Lemma 1.2 we have that \( T \Phi_n f = P_{T \mathcal{H}_n} f \to P_{R_T} f \) as \( n \to \infty \).

By assumption, \( \{ f_i \}_{i \in I} \) is a frame for \( N_T^\perp \). This implies that
\[
\Phi_n f \in \text{span}\{ f_i \}_{i \in I} \subseteq N_T^\perp
\]
for all \( n \). Since \( T^\dagger T \) is the orthogonal projection onto \( N_T^\perp \), we have
\[
\Phi_n f = T^\dagger T \Phi_n f \to T^\dagger P_{R_T} f = T^\dagger f \quad \text{as } n \to \infty.
\]

Remark. Observe that \( V_n \) is an operator on the finite-dimensional space \( T \mathcal{H}_n \). Therefore the coefficients \( \{ \langle f, V_n^{-1} T f_i \rangle \}_{i \in I_n} \) appearing in Theorem 1.4 can be calculated by finite-dimensional methods.

The reason for using frames instead of just bases in the approximation of \( T^\dagger \) is that frames are more flexible tools than bases in that they allow more properties to be satisfied simultaneously. Let us mention a few examples to illustrate this for some important frames for \( L^2(\mathbb{R}) \):

(i) A Gabor frame has the form \( \{ e^{2\pi ibx} g(x - na) \}_{m,n \in \mathbb{Z}} \), where \( g \in L^2(\mathbb{R}) \) and \( a, b > 0 \). By the Balian–Low Theorem, a function \( g \) generating a Gabor Riesz basis cannot be well-localized in both time and frequency; cf. [1]. But Gabor frames with this feature exist.

(ii) A wavelet frame has the form \( \{ a^{j/2} \psi(x - kb) \}_{j,k \in \mathbb{Z}} \), where \( \psi \in L^2(\mathbb{R}) \), \( a > 1 \), \( b > 0 \). \( L^2(\mathbb{R}) \) does not have an orthonormal basis of the form (7), for which the function \( \psi \) is \( C^\infty \) with exponential decay. However, frames of this type exist. A famous example is the Mexican hat \( \psi(x) = (1 - x^2)e^{-x^2/2} \); cf. [6].
2. NONLINEAR ITERATIVE APPROXIMATION OF $T^\dagger f$

In the previous section, the index sets $\{I_n\}_{n=1}^\infty$ were fixed independently of $f$. As a consequence, we obtained a family of operators $\{\Phi_n\}$ converging to $T^\dagger$ in the strong operator topology. We now describe an element-dependent method for approximation of $T^\dagger f$. This means that we fix $f \in \mathcal{H}$ and that the choice of $\{I_n\}_{n=1}^\infty$ depends on $f$. The advantage is that the choice of $I_n$ at the $n$th step of the approximation might fit $f$ better, but the disadvantage is that the method becomes nonlinear. This method is inspired by various versions of matching pursuit algorithms; cf. [8, 9].

Corresponding to an index set $I_n$, we use the notation $\mathcal{H}_n, S_n, V_n$ as in Section 1. First, fix $f \in \mathcal{H}$ and let $\epsilon > 0$ be given. Choose the set $I_1$ such that

$$\|P_{R_1}f - P_{T\mathcal{H}_1}P_{R_1}f\| \leq \frac{\epsilon}{2}.$$  

Write $P_{R_1}f = P_{T\mathcal{H}_1}P_{R_1}f + R_1$ and observe that

$$\|R_1\| = \|P_{R_1}f - P_{T\mathcal{H}_1}P_{R_1}f\| \leq \frac{\epsilon}{2},$$
$$\|P_{R_1}f\|^2 = \|P_{T\mathcal{H}_1}P_{R_1}f\|^2 + \|R_1\|^2.$$

Now choose $I_2$ such that

$$\|R_1 - P_{T\mathcal{H}_2}R_1\| \leq \frac{\epsilon}{4}.$$  

Write $R_1 = P_{T\mathcal{H}_2}R_1 + R_2$ and observe that

$$\|R_2\| = \|R_1 - P_{T\mathcal{H}_2}R_1\| \leq \frac{\epsilon}{4}, \quad \|R_1\|^2 = \|P_{T\mathcal{H}_2}R_1\|^2 + \|R_2\|^2.$$  

Thus

$$P_{R_2}f = P_{T\mathcal{H}_1}P_{R_1}f + P_{T\mathcal{H}_2}R_1 + R_2$$
$$\|P_{R_2}f\|^2 = \|P_{T\mathcal{H}_1}P_{R_1}f\|^2 + \|P_{T\mathcal{H}_2}R_1\|^2 + \|R_2\|^2$$
$$\|P_{R_2}f - (P_{T\mathcal{H}_1}P_{R_1}f + P_{T\mathcal{H}_2}R_1)\| = \|R_2\| \leq \frac{\epsilon}{4}.$$  

In general, after constructing $R_n$, choose $I_{n+1}$ such that

$$\|R_n - P_{T\mathcal{H}_{n+1}}R_n\| \leq \frac{\epsilon}{2^{n+1}}.$$  

Write $R_n = P_{T\mathcal{H}_{n+1}}R_n + R_{n+1}$ and observe that

$$\|R_{n+1}\| = \|R_n - P_{T\mathcal{H}_{n+1}}R_n\| \leq \frac{\epsilon}{2^{n+1}}, \quad \|R_n\|^2 = \|P_{T\mathcal{H}_{n+1}}R_n\|^2 + \|R_{n+1}\|^2.$$
Thus, with $R_0 := P_{R_T} f$ we have

$$P_{R_T} f = \sum_{k=0}^{n} P_{T \mathcal{H}_{k+1}} R_k + R_{n+1},$$

$$\left\| P_{R_T} f - \sum_{k=0}^{n} P_{T \mathcal{H}_{k+1}} R_k \right\| = \left\| R_{n+1} \right\| \leq \frac{\epsilon}{2n+1},$$

$$\left\| P_{R_T} f \right\|^2 = \sum_{k=0}^{n} \left\| P_{T \mathcal{H}_{k+1}} R_k \right\|^2 + \left\| R_{n+1} \right\|^2.$$ 

Since $\{T f_i\}_{i \in I_{k+1}}$ is a frame for $T \mathcal{H}_{k+1}$ and the corresponding frame operator is $V_{k+1}$, we have by Lemma 1.1 that

$$P_{T \mathcal{H}_{k+1}} R_k = \sum_{i \in I_{k+1}} \langle R_k, V_{k+1}^{-1} T f_i \rangle T f_i = T \sum_{i \in I_{k+1}} \langle R_k, V_{k+1}^{-1} T f_i \rangle f_i.$$

Thus $P_{T \mathcal{H}_{k+1}} R_k = T g_k$, where

$$g_k = \sum_{i \in I_{k+1}} \langle R_k, V_{k+1}^{-1} T f_i \rangle f_i.$$

Thus

$$\left\| P_{R_T} f - T \sum_{k=0}^{n} g_k \right\| \leq \frac{\epsilon}{2n+1}.$$

The iterative approximation of $P_{R_T} f$ leads to the following result on approximation of $T^\dagger f$.

**Theorem 2.1.** Fix $f \in \mathcal{H}$ and construct $\{g_k\}_{k=0}^\infty$ as above. Then

$$\left\| T^\dagger f - \sum_{k=0}^{n} g_k \right\| \leq \frac{\epsilon}{2n+1} \|T^\dagger\|.$$

**Proof.** First observe that each $g_k \in \text{span}\{f_i\}_{i \in I} \subseteq N_T^\perp$. Also, $T^\dagger f \in R_{T^\dagger} = N_T^\perp$, and since $T^\dagger T$ is the orthogonal projection onto $N_T^\perp$, we obtain

$$\left\| T^\dagger f - \sum_{k=0}^{n} g_k \right\| = \left\| T^\dagger T \left( T^\dagger f - \sum_{k=0}^{n} g_k \right) \right\| \leq \|T^\dagger\| \cdot \left\| T T^\dagger f - T \sum_{k=0}^{n} g_k \right\| \leq \|T^\dagger\| \cdot \left\| P_{R_T} f - T \sum_{k=0}^{n} g_k \right\| \leq \frac{\epsilon}{2n+1} \|T^\dagger\|.$$
Let $P$ denote the orthogonal projection onto $\text{span}\{Tf_i\}_{i \in I_k}$. Then

$$\|f - Pf\| \leq \left\| f - T \sum_{k=0}^{n} g_k \right\|.$$  

By writing $Pf = Tg$, $g$ might approximate $T^* f$ even better than $\sum_{k=0}^{n} g_k$. However, for large index sets, calculation of $g$ becomes more involved because of the need to invert the frame operator corresponding to $\{Tf_i\}_{i \in I_k}$.

The motivation behind the iterative method is to “split the inversion into successive inversions of smaller matrices.” However, in the worst case the index set $I_n$ may have a lot of overlap with $I_1, I_2, \ldots, I_{n-1}$ (or even include those sets) and then the iterative method is not appropriate.

3. APPROXIMATION OF THE INVERSE FRAME OPERATOR

In this section we consider a frame $\{g_i\}_{i \in I}$ for $\mathcal{H}$ with frame bounds $A, B$. Since the frame operator $S$ corresponding to $\{g_i\}_{i \in I}$ is defined on $\mathcal{H}$, which is usually infinite-dimensional, it is a non-trivial matter to invert $S$. Therefore it is natural to try to approximate $S^{-1}$; cf. [3, 4]. We now describe how the methods from the previous sections can be used to approximate $S^{-1}$.

First, we apply the method in Section 1 directly to the operator $T := S$. Since $S$ is invertible, the pseudo-inverse of $S$ equals $S^{-1}$. The construction in Section 1 is based on the choice of a frame $\{f_i\}_{i \in I}$ for $\mathcal{H}$; since $S$ is injective, we have $N_{\frac{1}{2}} = N_{\frac{1}{2}} = \mathcal{H}$, so we can choose $\{f_i\}_{i \in I} := \{g_i\}_{i \in I}$. Then, by Theorem 1.4,

$$\Phi_n f := \sum_{i \in I_n} \langle f, V_n^{-1} S f_i \rangle f_i \to S^{-1} f \quad \text{as } n \to \infty,$$

where

$$V_n : \mathcal{H}_n \to \mathcal{H}_n, \quad V_n = SS_n S_n.$$  

We could also apply the iterative method from Section 2. Since $\|S^{-1}\| \leq \frac{1}{A}$, Theorem 2.1 gives the estimate

$$\left\| S^{-1} f - \sum_{k=0}^{n} S_k \right\| \leq \frac{\epsilon}{A^{2n+1}}.$$  

In many contexts, the frame operator is only needed to calculate the frame coefficients $\{\langle f, S^{-1} g_i \rangle \}_{i \in I}$. In that case a different method can be used. Given the frame $\{g_i\}_{i \in I}$ we consider the operator

$$T : \ell^2(I) \to \mathcal{H}, \quad T\{c_i\}_{i \in I} = \sum_{i \in I} c_i g_i.$$  

It is well known that $T$ is a bounded surjective operator. In [5, Theorem 3.1] it is shown that
\[ T^\dagger f = \{ (f, S^{-1}g_i) \}_{i \in I}, \quad f \in \mathcal{H}. \]

Now choose a frame $\{ f_i \}_{i \in I}$ for $N_{T^*}$. By Theorem 1.4 it follows that
\[ \Phi_n f := \sum_{i \in I_n} \langle f, V_n^{-1} T f_i \rangle f_i \to \{ (f, S^{-1}g_i) \}_{i \in I}, \quad \forall f \in \mathcal{H}, \]
where
\[ V_n : T\mathcal{H} \to T\mathcal{H}, \quad V_n = T S_n P_n T^*. \]

ACKNOWLEDGMENTS

The author thanks the anonymous referee and the associate editor for several improvements on the presentation of the results.

REFERENCES