Several Important Theorems
by
Francis J. Narcowich
November, 2014

1 The Projection Theorem

Let \( \mathcal{H} \) be a Hilbert space. When \( V \) is a finite dimensional subspace of \( \mathcal{H} \) and \( f \in \mathcal{H} \), we can always find a unique \( p \in V \) such that \( \|f - p\| = \min_{v \in V} \|f - v\| \). This fact is the foundation of least-squares approximation. What happens when we allow \( V \) to be infinite dimensional? We will see that the minimization problem can be solved if and only if \( V \) is closed.

**Theorem 1.1** (The Projection Theorem). Let \( \mathcal{H} \) be a Hilbert space and let \( V \) be a subspace of \( \mathcal{H} \). For every \( f \in \mathcal{H} \) there is a unique \( p \in V \) such that \( \|f - p\| = \min_{v \in V} \|f - v\| \) if and only if \( V \) is a closed subspace of \( \mathcal{H} \).

To prove this, we need the following lemma.

**Lemma 1.2** (Polarization Identity). Let \( \mathcal{H} \) be a Hilbert space. For every pair \( f, g \in \mathcal{H} \), we have

\[
\|f + g\|^2 + \|f - g\|^2 = 2(\|f\|^2 + \|g\|^2).
\]

**Proof.** Adding the \( \pm \) identities \( \|f \pm g\|^2 = \|f\|^2 \pm \langle f, g \rangle \pm \langle g, f \rangle + \|g\|^2 \) yields the result. \( \square \)

The polarization identity is an easy consequence of having an inner product. It is surprising that if a norm satisfies the polarization identity, then the norm *comes* from an inner product.\(^1\)

**Proof.** (Projection Theorem) Showing that the existence of minimizer implies that \( V \) is closed is left as an exercise. So we assume that \( V \) is closed. For \( f \in \mathcal{H} \), let \( \alpha := \inf_{v \in V} \|v - f\| \). It is a little easier to work with this in an equivalent form, \( \alpha^2 = \inf_{v \in V} \|v - f\|^2 \). Thus, for every \( \varepsilon > 0 \) there is a \( v_\varepsilon \in V \) such that \( \alpha^2 \leq \|v_\varepsilon - f\|^2 < \alpha^2 + \varepsilon \). By choosing \( \varepsilon = 1/n \), where \( n \) is a positive integer, we can find a sequence \( \{v_n\}_{n=1}^{\infty} \) in \( V \) such that

\[
0 \leq \|v_n - f\|^2 - \alpha^2 < \frac{1}{n} \quad \text{(1.1)}
\]

Of course, the same inequality holds for a possibly different integer \( m \), \( 0 \leq \| v_m - f \|^2 - \alpha^2 < \frac{1}{m} \). Adding the two yields this:

\[
0 \leq \| v_n - f \|^2 + \| v_m - f \|^2 - 2\alpha^2 < \frac{1}{n} + \frac{1}{m}.
\]  

(1.2)

By polarization identity and a simple manipulation, we have

\[
\| v_n - v_m \|^2 + 4\| f - \frac{v_n + v_m}{2} \|^2 = 2(\| f - v_n \|^2 + \| f - v_m \|^2).
\]

We can subtract \( 4\alpha^2 \) from both sides and use (1.2) to get

\[
\| v_n - v_m \|^2 + 4(\| f - \frac{v_n + v_m}{2} \|^2 - \alpha^2) = 2(\| f - v_n \|^2 + \| f - v_m \|^2 - 2\alpha^2) < \frac{2}{n} + \frac{2}{m}.
\]

Because \( \frac{1}{2}(v_n + v_m) \in V \), \( \| f - \frac{v_n + v_m}{2} \|^2 \geq \inf_{v \in V} \| v - f \|^2 = \alpha^2 \). It follows that the second term on the left is nonnegative. Dropping it makes the left side smaller:

\[
\| v_n - v_m \|^2 < \frac{2}{n} + \frac{2}{m}.
\]  

(1.3)

As \( n, m \to \infty \), we see that \( \| v_n - v_m \| \to 0 \). Thus \( \{ v_n \}_{n=1}^\infty \) is a Cauchy sequence in \( \mathcal{H} \) and is therefore convergent to a vector \( p \in \mathcal{H} \). Since \( V \) is closed, \( p \in V \). Furthermore, taking limits in (1.1) implies that \( \| p - f \| = \inf_{v \in V} \| v - f \| \). The uniqueness of \( p \) is left as an exercise.

There are two important corollaries to this theorem; they follow from problem 4 of Assignment 1, 2021, and Theorem 1.1. We list them below.

**Corollary 1.3.** Let \( V \) be a subspace of \( \mathcal{H} \). There exists an orthogonal projection \( P : \mathcal{H} \to V \) for which \( \| f - Pf \| = \min_{v \in V} \| f - v \| \) if and only if \( V \) is closed.

**Corollary 1.4.** Let \( V \) be a closed subspace of \( \mathcal{H} \). Then, \( \mathcal{H} = V \oplus V^\perp \) and \( (V^\perp)^\perp = V \).

## 2 The Riesz Representation Theorem

Let \( V \) be a Banach space. A bounded linear transformation \( \Phi \) that maps \( V \) into \( \mathbb{R} \) or \( \mathbb{C} \) is called a **linear functional** on \( V \). The linear functionals form a Banach space \( V^* \), called the **dual space** of \( V \), with norm defined by

\[
\| \Phi \|_{V^*} := \sup_{v \neq 0} \frac{|\Phi(v)|}{\| v \|}.
\]
2.1 The linear functionals on Hilbert space

**Theorem 2.1 (The Riesz Representation Theorem).** Let $\mathcal{H}$ be a Hilbert space and let $\Phi : \mathcal{H} \to \mathbb{C}$ (or $\mathbb{R}$) be a bounded linear functional on $\mathcal{H}$. Then, there is a unique $g \in \mathcal{H}$ such that, for all $f \in \mathcal{H}$, $\Phi(f) = \langle f, g \rangle$.

**Proof.** The functional $\Phi$ is a bounded operator that maps $\mathcal{H}$ into the scalars. It follows from our discussion of bounded operators that the null space of $\Phi$, $N(\Phi)$, is closed. If $N(\Phi) = \mathcal{H}$, then $\Phi(f) = 0$ for all $f \in \mathcal{H}$, hence $\Phi = 0$. Thus we may take $g = 0$. If $N(\Phi) \neq \mathcal{H}$, then, since $N(\Phi)$ is closed, we have that $\mathcal{H} = N(\Phi) \oplus N(\Phi)\perp$. In addition, since $N(\Phi) \neq \mathcal{H}$, there exists a nonzero vector $h \in N(\Phi) \perp$. Moreover, $\Phi(h) \neq 0$, because $h$ is not in the null space $N(\Phi)$. Next, note that for $f \in \mathcal{H}$, the vector $w := \Phi(h)f - \Phi(f)h$ is in $N(\Phi)$. To see this, observe that

$$\Phi(w) = \Phi(\Phi(h)f - \Phi(f)h) = \Phi(h)\Phi(f) - \Phi(f)\Phi(h) = 0.$$ 

Because $w = \Phi(h)f - \Phi(f)h \in N(\Phi)$, it is orthogonal to $h \in N(\Phi)\perp$, we have that

$$0 = \langle \Phi(h)f - \Phi(f)h, h \rangle = \Phi(h)\langle f, h \rangle - \Phi(f)\langle h, h \rangle \underbrace{\text{.} \lVert h \rVert^2}_{\text{.}}.$$ 

Solving this equation for $\Phi(f)$ yields $\Phi(f) = \langle f, \frac{\Phi(h)}{\lVert h \rVert^2}h \rangle$. The vector $g := \frac{\Phi(h)}{\lVert h \rVert^2}h$ then satisfies $\Phi(f) = \langle f, g \rangle$. To show uniqueness, suppose $g_1, g_2 \in \mathcal{H}$ satisfy $\Phi(f) = \langle f, g_1 \rangle$ and $\Phi(f) = \langle f, g_2 \rangle$. Subtracting these two gives $\langle f, g_2 - g_1 \rangle = 0$ for all $f \in \mathcal{H}$. Letting $f = g_2 - g_1$ results in $\langle g_2 - g_1, g_2 - g_1 \rangle = 0$. Consequently, $g_2 = g_1$.

2.2 Adjoints of bounded linear operators

We now turn the problem of showing that an adjoint for a bounded operator always exists. This is just a corollary of the Riesz Representation Theorem.

**Corollary 2.2.** Let $L : \mathcal{H} \to \mathcal{H}$ be a bounded linear operator. Then there exists a bounded linear operator $L^* : \mathcal{H} \to \mathcal{H}$, called the adjoint of $L$, such that $\langle Lf, h \rangle = \langle f, L^* h \rangle$, for all $f, h \in \mathcal{H}$.

**Proof.** Fix $h \in \mathcal{H}$ and define the linear functional $\Phi_h(f) = \langle Lf, h \rangle$. Using the boundedness of $L$ and Schwarz’s inequality, we have $|\Phi_h(f)| \leq \|L\|\|f\|\|h\| = K\|f\|$, and so $\Phi_h$ is bounded. By Theorem 2.1 there is a
unique vector \( g \) in \( \mathcal{H} \) for which \( \Phi_h(f) = \langle f, g \rangle \). The vector \( g \) is uniquely determined by \( \Phi_h \); thus \( g = g_h \) a function of \( h \). We claim that \( g_h \) is a linear function of \( h \). Consider \( h = ah_1 + bh_2 \). Note that \( \Phi_h(f) = \langle Lf, ah_1 + bh_2 \rangle = \bar{a} \Phi_{h_1}(f) + b \Phi_{h_2}(f) \). Since \( \Phi_{h_1}(f) = \langle f, g_1 \rangle \) and \( \Phi_{h_2}(f) = \langle f, g_2 \rangle \), we see that

\[
\Phi_h(f) = \langle f, g_h \rangle = \bar{a} \Phi_{h_2}(f) + b \Phi_{h_2}(f) = \langle f, ag_{h_1} + bg_{h_2} \rangle.
\]

It follows that \( g_h = ag_{h_1} + bg_{h_2} \) and that \( g_h \) is a linear function of \( h \). It is also bounded. If \( f = g_h \), then \( \Phi_h(g_h) = \|g_h\|^2 \). From the bound \( |\Phi_h(f)| \leq \|L\| \|f\| \|h\| \), we have \( \|g_h\|^2 \leq \|L\| \|g_h\| \|h\| \). Dividing by \( \|g_h\| \) then yields \( \|g_h\| \leq \|L\| \|h\| \). Thus the correspondence \( h \rightarrow g_h \) is a bounded linear function on \( \mathcal{H} \). Denote this function by \( L^* \). Since \( \langle Lf, h \rangle = \langle f, g_h \rangle \), we have that \( \langle Lf, h \rangle = \langle f, L^*h \rangle \).

**Corollary 2.3.** \( \|L^*\| = \|L\| \).

*Proof.* By problem 7 in Assignment 7, 2021, \( \|L\| = \sup_{f,h} \|\langle Lf, h \rangle\| \), where \( \|h\| = \|f\| = 1 \). On the other hand, \( \|L^*\| = \sup_{f,h} \|\langle L^*h, f \rangle\| \). Since \( \langle L^*h, f \rangle = \langle f, L^*h \rangle \), we have that \( \sup_{f,h} \|\langle L^*h, f \rangle\| = \sup_{f,h} \|\langle Lf, h \rangle\| \). It immediately follows that \( \|L^*\| = \|L\| \).

**Example 2.4.** Let \( R = [0, 1] \times [0, 1] \) and suppose that \( k \) is a Hilbert-Schmidt kernel. If \( Lu(x) = \int_0^1 k(x, y)u(y)dy \), then \( L^*v(x) = \int_0^1 \bar{k}(y, x)v(y)dy \).

*Proof.* We will use \( s, t \) as the integration variables and switch back, to avoid confusion. We begin with \( \langle Lu, v \rangle = \int_0^1 (\int_0^1 k(s, t)u(t)dt)v(s)ds \). By Fubini’s theorem, we may switch the variables of integration to get this:

\[
\int_0^1 \left( \int_0^1 k(s, t)u(t)dt \right)v(s)ds = \int_0^1 \left( \int_0^1 k(s, t)v(s)ds \right)u(t)dt
= \int_0^1 \left( \int_0^1 \bar{k}(s, t)v(s)ds \right)u(t)dt.
\]

The result follows by changing variables from \( t, s \) to \( x, y \) in the second equation above.

**3 The Fredholm Alternative**

**Theorem 3.1** (The Fredholm Alternative). Let \( L : \mathcal{H} \to \mathcal{H} \) be a bounded linear operator whose range, \( R(L) \), is closed. Then, the equation \(Lf = g\)
and be solved if and only if \( \langle g, v \rangle = 0 \) for all \( v \in N(L^*) \). Equivalently, \( R(L) = N(L^*)^\perp \).

**Proof.** Let \( g \in R(L) \), so that there is an \( h \in \mathcal{H} \) such that \( g = Lh \). If \( v \in N(L^*) \), then \( \langle g, v \rangle = \langle Lh, v \rangle = \langle h, L^*v \rangle = 0 \). Consequently, \( R(L) \subseteq N(L^*)^\perp \). Let \( f \in N(L^*)^\perp \). Since \( R(L) \) is closed, the projection theorem, Theorem 1.1 and Corollary 1.3 imply that there exists an orthogonal projection \( P \) onto \( R(L) \) such that \( Pf \in R(L) \) and \( f' = f - Pf \in R(L)^\perp \). Moreover, since \( f \) and \( Pf \) are both in \( N(L^*)^\perp \), we have that \( f' \in R(L)^\perp \cap N(L^*)^\perp \). Hence, \( \langle Lh, f' \rangle = 0 = \langle h, L^*f' \rangle \), for all \( h \in \mathcal{H} \). Setting \( h = L^*f' \) then yields \( L^*f' = 0 \), so \( f' \in N(L^*) \). But \( f' \in N(L^*)^\perp \) and is thus orthogonal to itself; hence, \( f' = 0 \) and \( f = Pf \in R(L) \). It immediately follows that \( N(L^*)^\perp \subseteq R(L) \). Since we already know that \( R(L) \subseteq N(L^*)^\perp \), we have \( R(L) = N(L^*)^\perp \).

We want to point out that \( R(L) \) being closed is crucial for the theorem to be true. If it is not closed, then the projection \( P \) will not exist and the proof breaks down. In that case, one actually has \( R(L) = N(L^*)^\perp \), but not \( R(L) = N(L^*)^\perp \).

The theorem is stated in a variety of ways. The form that emphasizes the “alternative” is given in the result below, which follows immediately from Theorem 3.1.

**Corollary 3.2.** Let \( L : \mathcal{H} \to \mathcal{H} \) be a bounded linear operator whose range, \( R(L) \), is closed. Then, either the equation \( Lf = g \) has a solution or there exists a vector \( v \in N(L^*) \) such that \( \langle g, v \rangle \neq 0 \).