Math 7330: Functional Analysis Fall 2002
Notes A. Sengupta

In the following H is a complex Hilbert space.

1 Orthogonal Projections

We shall study orthogonal projections onto closed subspaces of H. In summary, we show:

- If X is any closed subspace of H then there is a bounded linear operator $P : H \to H$ such that P $= X$ and each element x can be written unqiuely as a sum $a + b$, with $a \in \mathrm{Im}(P)$ and $b \in ker(P)$; explicitly, $a = Px$ and $b = x - Px$. The point Px is the point on X closest to x. The operator P satisfies $P^2 = P$ and is self-adjoint.
- Conversely, if P_1 is any bounded linear operator $H \to H$ for which $P_1^2 = P_1$ then the following are equivalent: (i) P_1 is an orthogonal projection onto a closed subspace, (ii) P_1 is self-adjoint, (iii) P_1 is normal, i.e. commutes with its adjoint P_1^*.

1.1 Point in a convex set closest to a given point

Let C be a closed convex subset of H. We will prove that

there is a unique point in C which is closest to the origin.

This will use convexity of C, that C is closed, the fact that the topology on H arises from an inner-product, and that this topology makes H a complete metric space.

Let

$$r = d(0, C) = \inf\{|x| : x \in C\}$$

The function

$$f : H \to \mathbf{R} : x \mapsto d(x, C) = \inf\{|x - y| : y \in C\}$$

is continuous because it satisfies

$$|f(x) - f(y)| \le d(x, y) \tag{1}$$

This inequality can be verified as follows: for any for any $a \in C$ the triangle inequality implies

$$f(x) \le d(x, a) \le d(x, y) + d(y, a)$$

and so taking inf over $a \in C$ gives $f(x) \le d(x, y) + f(y)$; this implies $f(x) - f(y) \le d(x, y)$, and then we can interchange x and y.

Since C is closed, there is a sequence of points $x_n \in C$ with $|x_n| \to r$ as $n \to \infty$. We will show that the sequence (x_n) is necessarily Cauchy. This uses the parallelogram property:

$$|x_n - x_m|^2 + |x_n + x_m|^2 = 2(|x_n|^2 + |x_m|^2)$$

which gives

$$|x_n - x_m|^2 = 2(|x_n|^2 + |x_m|^2) - 4|\frac{x_n + x_m}{2}|^2$$

Since C is convex, the midpoint $\frac{x_n+x_m}{2}$ lies in C. Since r is the closes distance of C from 0, it follows that $|\frac{x_n+x_m}{2}| \geq r$ and so

$$|x_n - x_m|^2 \leq 2(|x_n|^2 + |x_m|^2) - 4r^2$$

If we let $n, m \to \infty$ the the right side approaches $2(r^2 + r^2) - 4r^2 = 0$, and so the sequence $(x_n)_n$ is Cauchy. Since H is complete, this sequence has a limit, say $\lim_{n\to\infty} x_n = a$. The continuity of the function f above implies that $d(a, C) = \lim_{n\to\infty} d(x_n, C) = r$. Thus *$a$ is a point in C closest to the origin.* If b is also another such point then the parallelogram property gives:

$$|a - b|^2 \leq 2(|a|^2 + |b|^2) - 4r^2 = 2(r^2 + r^2) - 4r^2 = 0$$

which shows that the point a is *unique*.

Now with C as above consider any point $x \in H$. The translate $x - C = \{x - c : c \in C\}$ is also a closed convex set and so there is a unique point c in C for which $|x - c|$ is smallest. Thus

Proposition 1 *If C is any non-empty closed convex subset of a Hilbert space H then there is a unique point in C closest to any given point x of H.*

1.2 The orthogonal projection on a closed subspace

Now let X be a *closed subspace* of H ('subspace' here means a linear subspace). So X is a closed convex set.

Let x be any point of H. Then there is a unique point in X closest to x. Denote this point by Px.

We shall prove that *$x - Px$ is orthogonal to X*.

Consider any $y \in X$. We will show that $(x - Px, y)$ is 0. If y is 0 this is clear; so lets normalize and assume $|y| = 1$. Note that everything takes place now in the three-dimensional space spanned by x, Px, y and basically we are saying that $x - Px$ is orthogonal to the subspace spanned by x and y. Let $\alpha \in \mathbf{C}$; then

$$|x - (Px + y)|^2 = |(x - Px) - y|^2 = |x - Px|^2 - 2\mathrm{Re}[\alpha(x - Px, y)] + |\alpha|^2|y|^2$$

If we make the choice $\alpha = (y, x - Px)$ the right side becomes

$$|x - Px|^2 - 2|(y, x - Px)|^2 + |(y, x - Px)|^2|y|^2 = |x - Px|^2 - |(y, x - Px)|^2$$

Thus the distance of x from the point $Px + y \in X$ would be less than $|x - Px|$ unless $(y, x - Px)$ is 0. This completes the argument.

Finally we show that $P : H \to X : x \mapsto Px$ is a linear map.

Let us write

$$Qx = x - Px$$

for all $x \in X$. We have shown above that Qx is orthogonal to X. Let $x, y \in X$ and $\alpha, \beta \in \mathbf{C}$. Then we have

$$\begin{aligned} P(\alpha x + \beta y) + Q(\alpha x + \beta y) &= \alpha x + \beta y \\ &= \alpha(Px + Qx) + \beta(Py + Qy) \\ &= (\alpha Px + \beta Py) + (\alpha Qx + \beta Qy) \end{aligned}$$

Moving the P terms to the left and the Q terms to the right we get

$$P(\alpha x + \beta y) - (\alpha Px + \beta Py) = (\alpha Qx + \beta Qy) - Q(\alpha x + \beta y)$$

The left side is in X and the right side is orthogonal to X. Therefore both sides must be 0. This implies that P and Q are both *linear*.

In summary:

Theorem 2 . *Let X be a closed subspace of a Hilbert space H. For each $x \in H$ there is a unique point Px in X closes to x. The mapping $P : H \to H : x \mapsto Px$ is linear. For any $x \in X$ there is a unique $a \in X$ and $b \in X^\perp$ such that $x = a + b$. In fact, $a = Px$ and $b = x - Px$. Thus*

$$H = X \oplus X^\perp$$

The only thing we didn't check is the uniqueness of the decomposition. But if $x = a + b$ with $a \in X$ and $b \in X^\perp$ then writing $x = Px + Qx$ we have $a - Px = Qx - b$, the left side being in X and the right side in $X^\perp$ we conclude that $a = Px$ and $b = Qx$.

The map P is called the *orthogonal projection* onto the closed subspace X.

Note that for any $x \in H$ we have

$$|x|^2 = |Px|^2 + |Qx|^2$$

which implies, in particular, that P is a bounded linear map.

Proposition 3 *Let X be a closed subspace of X and $j : X \to H : x \mapsto x$ the inclusion map. The the adjoint $j^* : H \to X$ is given by $j^*x = Px$ for every $x \in X$.*

Proof. For any $a \in H$ and $x \in X$:

$$\begin{aligned} (j^*a, x) &= (a, jx) \\ &= (a, x) \\ &= (Pa, x) \end{aligned}$$

Since Pa, j^*a are in X and the above holds for all $x \in X$ it follows that $Pa = j^*a$ for all $a \in H$. $\boxed{\text{QED}}$

1.3 Projection operators

Let V be a vector space. A map $A : V \to V$ is a *projection operator* if it is linear and satisfies

$$A^2 = A$$

In this subsection we shall assume that $A : V \to V$ is a projection operator.

Observe that $I - A$ is also then a projection operator:

$$(I - A)^2 = I - 2A + A^2 = I - A$$

If a point y lies in the image of A then it is of the form Ax, for some $x \in V$, and so then $Ay = A(Ax) = A^2x = Ax = y$; thus

$$Ay = y \text{ if and only if } y \text{ is in the image of } A \tag{2}$$

Put another way,

$$\mathrm{Im}(A) = \ker(I - A) \tag{3}$$

Applying this result to the projection operator $I - A$ gives

$$\ker(A) = \mathrm{Im}(I - A) \tag{4}$$

Any vector $x \in V$ can be expressed as

$$x = Ax + (I - A)x$$

where the first term Ax is clearly in the image of A while the second term is in $\ker(A)$. Furthermore, this decomposition is unique since any element y which is in both $\ker(A)$ and $\mathrm{Im}(A)$ must be 0 because $y \in \mathrm{Im}(A)$ implies $y = Ay$ while $y \in \ker(A)$ means $Ay = 0$.

Thus V splits into a direct sum of the subspace $\mathrm{Im}(A)$ and $\ker(A)$:

$$V = \mathrm{Im}(A) \oplus \ker(A) \tag{5}$$

1.4 Characterization of orthogonal projections

We have shown in class that P is self-adjoint and satisfies $P^2 = P$.

We have also seen in class that for a bounded linear map $P : H \to H$ for which $P^2 = P$ the following are equivalent: (a) P is normal, (b) P is self-adjoint, (c) P is an orthogonal projection.

First let us prove a couple of useful facts:

Lemma 4 *Let H be a complex vector space with a Hermitian inner-product $(\cdot, \cdot)$. Let $A : H \to H$ be a bounded linear map. Then:*

(i) *If $(Ax, x) = 0$ for all $x \in H$ then $A = 0$*

(ii) *the operator A is normal, i.e. satisfies $AA^* = A^*A$, if and only if $|Ax| = |A^*x|$ for every $x \in H$. In particular, if A is normal then $\ker(A) = \ker(A^*)$.*

Proof. (i) Suppose $(Ax, x) = 0$ for all $x \in H$. Replacing x by $x + y$ we get

$$(Ax, y) + (Ay, x) = 0 \tag{6}$$

In this replace y by iy to get

$$-i(Ax, y) + i(Ay, x) = 0$$

which says

$$(Ax, y) - (Ay, x) = 0 \tag{7}$$

Combining (6) and (7) we get

$$(Ax, y) = 0$$

for all $x, y \in H$. Taking $y = Ax$ shows that $|Ax|^2 = 0$ for all $x \in H$, so $Ax = 0$ for all $x \in H$, i.e. $A = 0$.

The proof of (ii) now follows from:

$$\big((AA^* - A^*A)x, x\big) = (A^*x, A^*x) - (Ax, Ax) = |A^*x|^2 - |Ax|^2$$

which shows that $|Ax| = |A^*x|$ for all $x \in H$ if and only if $\big((AA^* - A^*A)x, x\big)$ is 0 for all $x \in H$. $\boxed{\text{QED}}$

Here is a useful characterization of orthogonal projections:

Proposition 5 *Let $P : H \to H$ be a bounded linear map on the complex Hilbert space H such that $P^2 = P$. Then the following are equivalent:*

(i) *P is self-adjoint*

(ii) *P is normal*

(iii) *$x - Px$ is orthogonal to Px for every $x \in H$.*

If these conditions hold then P is the orthogonal projection onto its image.

Proof. If P is self-adjoint then of course P is normal.

Now suppose P is a normal operator which is a projection, i.e. satisfies $P^2 = P$. Then:

$$\begin{aligned}
\mathrm{Im}(P)^\perp &= \ker(P^*) && \text{true for any bounded linear operator} \\
&= \ker(P) && \text{because } P \text{ is normal} \\
&= \mathrm{Im}(I - P) && \text{because } P \text{ is a projection, by (4)}
\end{aligned}$$

In particular, $(I - P)x \in \mathrm{Im}(I - P)$ is orthogonal to $Px \in \mathrm{Im}(P)$.

So now suppose $P : H \to H$ is a projection operator for which $x - Px$ is orthgonal to Px for all $x \in H$. Then

$$\big(x, (P - P^*P)x\big) = \big((I - P)x, Px\big) = 0$$

for all $x \in H$. Since H is a *complex* Hilbert space, it follows that

$$P = P^* P$$

Taking the adjoint of this equation gives

$$P^* = P^* P$$

So

$$P = P^*$$

Thus P is self-adjoint.

Assume P satisfies the conditions above. Then $\mathrm{Im}(P) = \ker(I - P)$ is a *closed* subspace. For every $x \in H$ we have the decomposition $x = Px + (x - Px)$ with $Px \in \mathrm{Im}(P)$ and $x - Px$ is orthogonal to $\mathrm{Im}(P)$, as shown above. This means that P is the orthogonal projection onto the closed subspace $\mathrm{Im}(P)$. $\boxed{\text{QED}}$