Introduction to Empirical Processes and Semiparametric Inference Lecture 22: Preliminaries for Semiparametric Inference

Michael R. Kosorok, Ph.D.

Professor and Chair of Biostatistics

Professor of Statistics and Operations Research

University of North Carolina-Chapel Hill

Projections

Geometrically, the projection of an object T onto a space $\mathcal{S}$ is the element $\hat{S} \in \mathcal{S}$ that is "closest" to T, provided such an element exists.

In the semiparametric inference context,

- the object is usually a random variable and
- the spaces of interest for projection are usually sets of square-integrable random variables.

The following theorem gives a simple method for identifying the projection in this setting:

Theorem 17.1 *Let $\mathcal{S}$ be a linear space of real random variables with finite second moments.*

Then $\hat{S}$ is the projection of T onto $\mathcal{S}$ if and only if

(i) $\hat{S} \in S$ and

(ii) $E(T - \hat{S})S = 0$ for all $S \in \mathcal{S}$.

If S_1 *and* S_2 *are both projections, then* $S_1 = S_2$ *almost surely.*

If $\mathcal{S}$ *contains the constants, then* $ET = E\hat{S}$ *and*

$$\text{cov}(T - \hat{S}, S) = 0$$

for all $S \in \mathcal{S}$.

Proof. First assume (i) and (ii) hold.

Then, for any $S \in \mathcal{S}$, we have

$$E(T-S)^2 \quad = \quad E(T-\hat{S})^2 + 2E(T-\hat{S})(\hat{S}-S) + E(\hat{S}-S)^2. \quad (1)$$

But Conditions (i) and (ii) force the middle term to be zero, and thus $E(T-S)^2 \geq E(T-\hat{S})^2$ with strict inequality whenever $E(\hat{S}-S)^2 > 0.$

Thus $\hat{S}$ is the almost surely unique projection of T onto $\mathcal{S}$.

Conversely, assume that $\hat{S}$ is a projection and note that for any $\alpha \in \mathbb{R}$ and any $S \in \mathcal{S}$,

$$E(T - \hat{S} - \alpha S)^2 - E(T - \hat{S})^2 = -2\alpha E(T - \hat{S})S + \alpha^2 ES^2.$$

Since $\hat{S}$ is a projection, the left side is strictly nonnegative for every α.

But the parabola

$$\alpha \mapsto \alpha^2 ES^2 - 2\alpha E(T - \hat{S})S$$

is nonnegative for all α and S only if $E(T - \hat{S})S = 0$ for all S.

Thus (ii) holds, and (i) is part of the definition of a projection and so holds automatically.

The uniqueness follows from application of (1) to both S_1 and S_2, forcing $E(S_1 - S_2)^2 = 0$.

If the constants are in $\mathcal{S}$, then Condition (ii) implies that $E(T - \hat{S})c = 0$ for $c = 1$, and the theorem follows.$\Box$

Note that the theorem does not imply that a projection always exists.

In fact, if the set $\mathcal{S}$ is open in the $L_2(P)$ norm, then the infimum of $E(T-S)^2$ over $S \in \mathcal{S}$ is not achieved.

A sufficient condition for existence, then, is that $\mathcal{S}$ be closed in the $L_2(P)$ norm, but often existence can be established directly: we will discuss this more later.

A very useful example of a projection is a conditional expectation.

For X and Y real random variables on a probability space, $g_0(y) \equiv E(X|Y = y)$ is the conditional expectation of X given $Y = y$.

If we let $\mathcal{G}$ be the space of all measurable functions g of Y such that $Eg^2(Y) < \infty$, then it is easy to verify that

$$E(X - g_0(Y))g(Y) = 0, \quad \text{for every } g \in \mathcal{G}.$$

Thus, provided $Eg_0^2(Y) < \infty$, $E(X|Y = y)$ is the projection of X onto the linear space $\mathcal{G}$.

By Theorem 17.1, the conditional expectation is almost surely unique.

We will utilize conditional expectations frequently for calculating projections needed in semiparametric inference settings.

Hilbert Spaces

A Hilbert space is essentially an abstract generalization of a finite-dimensional Euclidean space.

This abstraction is a special case of a Banach space and, like a Banach space, is often infinite-dimensional.

To be precise, a Hilbert space is a Banach space with an inner product.

An inner product on a Banach space $\mathbb{D}$ with norm $\|\cdot\|$ is a function

$$\langle\cdot,\cdot\rangle : \mathbb{D} \times \mathbb{D} \mapsto \mathbb{R}$$

such that, for all $\alpha, \beta \in \mathbb{R}$ and $x, y, z \in \mathbb{D}$, the following hold:

(i) $\langle x, x\rangle = \|x\|^2$,

(ii) $\langle x, y\rangle = \langle y, x\rangle$, and

(iii) $\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$.

A semi-inner product arises when $\|\cdot\|$ is a semi-norm.

It is also possible to generate a norm beginning with an inner product.

Let $\mathbb{D}$ be a linear space with semi-inner product $\langle \cdot, \cdot \rangle$ (not necessarily with a norm), i.e., $\langle \cdot, \cdot \rangle$ satisfies Conditions (ii)–(iii) above and also satisfies $\langle x, x \rangle \geq 0$.

Then $\|x\| \equiv \langle x, x \rangle^{1/2}$, for all $x \in \mathbb{D}$, defines a semi-norm.

This is verified in the following theorem:

Theorem 17.2 *Let $\langle \cdot, \cdot \rangle$ be a semi-inner product on $\mathbb{D}$, with $\|x\| \equiv \langle x, x \rangle^{1/2}$ for all $x \in \mathbb{D}$.*

Then, for all $\alpha \in \mathbb{R}$ and all $x, y \in \mathbb{D}$,

(a) $\langle x, y \rangle \leq \|x\| \, \|y\|$,

(b) $\|x + y\| \leq \|x\| + \|y\|$, *and*

(c) $\|\alpha x\| = |\alpha| \times \|x\|$.

Moreover, if $\langle \cdot, \cdot \rangle$ is also an inner product, then $\|x\| = 0$ if and only if $x = 0$.

Part (a) in Theorem 17.2 is also known as the Cauchy-Schwartz inequality.

Two elements x, y in a Hilbert space $\mathbb{H}$ with inner product $\langle \cdot, \cdot \rangle$ are *orthogonal* if $\langle x, y \rangle = 0$, denoted $x \perp y$.

For any set $C \subset \mathbb{H}$ and any $x \in \mathbb{H}$, x is orthogonal to C if $x \perp y$ for every $y \in C$, denoted $x \perp C$.

The two subsets $C_1, C_2 \subset \mathbb{H}$ are orthogonal if $x \perp y$ for every $x \in C_1$ and $y \in C_2$, denoted $C_1 \perp C_2$.

For any $C_1 \subset \mathbb{H}$, the *orthocomplement* of C_1, denoted $C_1^{\perp}$, is the set $\{x \in \mathbb{H} : x \perp C_1\}$.

Let the subspace $H \subset \mathbb{H}$ be linear and closed.

By Theorem 17.1, we have for any $x \in \mathbb{H}$ that there exists an element $y \in H$ that satisfies $\|x - y\| \leq \|x - z\|$ for any $z \in H$, and such that $\langle x - y, z \rangle = 0$ for all $z \in H$.

Let Π denote an operator that performs this projection, i.e., let $\Pi x \equiv y$, where y is the projection of x onto H.

The "projection" operator $\Pi : \mathbb{H} \mapsto H$ has several important properties:

Theorem 17.3 *Let H be a closed, linear subspace of $\mathbb{H}$, and let $\Pi : \mathbb{H} \mapsto H$ be the projection operator onto H.*

Then

(i) Π *is continuous and linear,*

(ii) $\|\Pi x\| \leq \|x\|$ *for all* $x \in \mathbb{H}$,

(iii) $\Pi^2 \equiv \Pi\,\Pi = \Pi$, *and*

(iv) $N(\Pi) = H^{\perp}$ *and* $R(\Pi) = H$.

Proof. Let $x, y \in \mathbb{H}$ and $\alpha, \beta \in \mathbb{R}$.

If $z \in H$, then

$$\begin{aligned} \langle [\alpha x + \beta y] - [\alpha \Pi x + \beta \Pi y], z \rangle &= \alpha \langle x - \Pi x, z \rangle + \beta \langle y - \Pi y, z \rangle \\ &= 0. \end{aligned}$$

Now by the uniqueness of projections (Theorem 17.1), we now have that $\Pi(\alpha x + \beta y) = \alpha \Pi x + \beta \Pi y$.

This yields linearity of Π.

If we can establish (ii), (i) will follow.

Since $\langle x - \Pi x, \Pi x\rangle = 0$, for any $x \in \mathbb{H}$, we have

$$\|x\|^2 = \|x - \Pi x\|^2 + \|\Pi x\|^2 \geq \|\Pi x\|^2,$$

and (ii) (and hence also (i)) follows.

For any $y \in H$, $\Pi y = y$.

Thus for any $x \in \mathbb{H}$, $\Pi(\Pi x) = \Pi x$, and (iii) follows.

Now let $x \in N(\Pi)$.

Then $x = x - \Pi x \in H^{\perp}$.

Conversely, for any $x \in H^{\perp}$, $\Pi x = 0$ by definition, and thus $x \in N(\Pi)$.

Hence $N(\Pi) = H^{\perp}$.

Now it is trivial to verify that $R(\Pi) \subset H$ by the definitions.

Moreover, for any $x \in H$, $\Pi x = x$, and thus $H \subset R(\Pi)$.

This implies (iv).$\square$

One other point we note is that for any projection Π onto a closed linear subspace $H \subset \mathbb{H}$, $I - \Pi$, where I is the identity, is also a projection onto the closed linear subspace $H^{\perp}$.

A key example of a Hilbert space $\mathbb{H}$ is $\mathbb{H} = L_2(P)$ with inner product $\langle f, g \rangle = \int f g dP$.

A closed, linear subspace of interest to us is $L_2^0(P) \subset L_2(P)$ which consists of all mean zero functions in $L_2(P)$.

The projection operator $\Pi : L_2(P) \mapsto L_2^0(P)$ is $\Pi x = x - Px$.

To see this, note that $\Pi x \in L_2^0(P)$ and
$\langle x - \Pi x, y \rangle = \langle Px, y \rangle = Px\, Py = 0$ for all $y \in L_2^0(P)$.

Thus by Theorem 17.1, Πx is the unique projection onto $L_2^0(P)$.

It is also not hard to verify that $I - \Pi$ is the projection onto the constants (see Exercise 17.4.2).

We close this section with a brief discussion of *linear functionals* on Hilbert spaces.

Recall from Chapter 6 the definition of a linear operator and the fact that the norm for a linear operator $T : \mathbb{D} \mapsto \mathbb{E}$ is

$$\|T\| \equiv \sup_{x \in \mathbb{D} : \|x\| \leq 1} \|Tx\|.$$

In the special case where $\mathbb{E} = \mathbb{R}$, a linear operator is called a linear functional.

A linear functional T, like a linear operator, is bounded when $\|T\| < \infty$.

By Proposition 6.15, boundedness is equivalent to continuity in this setting.

We now present a very important result for bounded linear functionals in Hilbert spaces.

Theorem 17.5 (**Riesz representation theorem**) If $L : \mathbb{H} \mapsto \mathbb{R}$ is a bounded linear functional on a Hilbert space, then there exists a unique element $h_0 \in \mathbb{H}$ such that $L(h) = \langle h, h_0 \rangle$ for all $h \in \mathbb{H}$, and, moreover, $\|L\| = \|h_0\|$.

More on Banach Spaces

As with Hilbert spaces, a linear functional on a Banach space is just a linear operator with real range.

The *dual space* $\mathbb{B}^*$ of a Banach space $\mathbb{B}$ is the set of all continuous, linear functionals on $\mathbb{B}$.

By applying Proposition 6.15, it is clear that every $b^* \in \mathbb{B}^*$ satisfies $|b^*b| \leq \|b^*\| \, \|b\|$ for every $b \in \mathbb{B}$, where

$$\|b^*\| \equiv \sup_{b \in \mathbb{B}: \|b\| \leq 1} |b^*b| < \infty.$$

For the special case of a Hilbert space $\mathbb{H}$, $\mathbb{H}^*$ can identified with $\mathbb{H}$ by the Reisz representation theorem.

This implies that there exists an *isometry* (a one-to-one map that preserves norms) between $\mathbb{H}$ and $\mathbb{H}^*$.

To see this, choose $h^* \in \mathbb{H}^*$ and let $\tilde{h} \in \mathbb{H}$ be the unique element that satisfies $\langle h, \tilde{h} \rangle = h^* h$ for all $h \in \mathbb{H}$.

Then

$$\|h^*\| = \sup_{h \in \mathbb{H}: \|h\| \leq 1} |\langle h, \tilde{h} \rangle| \leq \|\tilde{h}\|$$

by the Cauchy-Schwartz inequality.

The desired conclusion follows since h^* was arbitrary.

We now return to the generality of Banach spaces.

For each continuous, linear operator between Banach spaces $A : \mathbb{B}_1 \mapsto \mathbb{B}_2$, there exists an *adjoint map* (or just adjoint) $A^* : \mathbb{B}_2^* \mapsto \mathbb{B}_1^*$ defined by

$$(A^* b_2^*) b_1 = b_2^* A b_1$$

for all $b_1 \in \mathbb{B}_1$ and $b_2^* \in \mathbb{B}_2^*$.

It is straightforward to verify that the resulting A^* is linear.

The following proposition tells us that A^* is also continuous (by being bounded):

Proposition 17.6 *Let* $A : \mathbb{B}_1 \mapsto \mathbb{B}_2$ *be a bounded linear operator between Banach spaces.*

Then $\|A^*\| = \|A\|$.

Proof. Since also, for any $b_2^* \in \mathbb{B}_2^*$,

$$\begin{aligned} \|A^* b_2^*\| &= \sup_{b_1 \in \mathbb{B}_1 : \|b_1\| \leq 1} |A^* b_2^* b_1| \\ &= \sup_{b_1 \in \mathbb{B}_1 : \|b_1\| \leq 1} \left\{ \left| b_2^* \left(\frac{A b_1}{\|A b_1\|} \right) \right| \|A b_1\| \right\} \\ &\leq \|b_2^*\| \, \|A\|, \end{aligned}$$

we have $\|A^*\| \leq \|A\|$.

Thus $\|A^*\|$ is a continuous, linear operator.

Now let $A^{**} : \mathbb{B}_1^{**} \mapsto \mathbb{B}_2^{**}$ be the adjoint of A^* with respect to the double duals (duals of the duals) of $\mathbb{B}_1$ and $\mathbb{B}_2$.

Note that for $j = 1, 2$, $\mathbb{B}_j \subset \mathbb{B}_j^{**}$, since for any $b_j \in \mathbb{B}_j$, the map $b_j : \mathbb{B}_j^* \mapsto \mathbb{R}$ defined by $b_j^* \mapsto b_j^* b_j$, is a bounded linear functional.

By the definitions involved, we now have for any $b_1 \in \mathbb{B}_1$ and $b_2^* \in \mathbb{B}_2^*$ that

$$(A^{**}b_1)b_2^* = (A^*b_2^*)b_1 = b_2^*Ab_1,$$

and thus $\|A^{**}\| \leq \|A^*\|$ and the restriction of A^{**} to $\mathbb{B}_1$, denoted hereafter A_1^{**}, equals A.

Hence $\|A\| = \|A_1^{**}\| \leq \|A^*\|$, and the desired result follows.$\square$

We can readily see that the adjoint of an operator $A : \mathbb{H}_1 \mapsto \mathbb{H}_2$ between two Hilbert spaces, with respective inner products $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$, is a map $A^* : \mathbb{H}_2 \mapsto \mathbb{H}_1$ satisfying

$$\langle Ah_1, h_2 \rangle_2 = \langle h_1, A^* h_2 \rangle_1$$

for every $h_1 \in \mathbb{H}_1$ and $h_2 \in \mathbb{H}_2$.

Here we are using the isometry for Hilbert spaces described above.

Now consider the adjoint of a restriction of a continuous linear operator $A : \mathbb{H}_1 \mapsto \mathbb{H}_2$, $A_0 : \mathbb{H}_{0,1} \subset \mathbb{H}_1 \mapsto \mathbb{H}_2$, where $\mathbb{H}_{0,1}$ is a closed, linear subspace of $\mathbb{H}_1$.

If $\Pi : \mathbb{H}_1 \mapsto \mathbb{H}_{0,1}$ is the projection onto the subspace, it is not hard to verify that the adjoint of A_0 is $A_0^* \equiv \Pi \circ A^*$ (see Exercise 17.4.3).

Recall from Chapter 6 the notation $B(\mathbb{D}, \mathbb{E})$ denoting the collection of all linear operators between normed spaces $\mathbb{D}$ and $\mathbb{E}$.

From Lemma 6.16, we know that for a given $T \in B(\mathbb{B}_1, \mathbb{B}_2)$, for Banach spaces $\mathbb{B}_1$ and $\mathbb{B}_2$, $R(T)$ is not closed unless T is continuously invertible.

We now give an illustrative counter-example.

Let $\mathbb{B}_1 = \mathbb{B}_2 = L_2(0,1)$, and define $T : L_2(0,1) \mapsto L_2(0,1)$ by $Tx(u) \equiv ux(u)$.

Then $\|T\| \leq 1$, and thus $T \in B(L_2(0,1), L_2(0,1))$.

However, it is clear that

$$R(T) \quad = \quad \left\{ y \in L_2(0,1) : \int_0^1 u^{-2} y^2(u) du < \infty \right\}.$$

Although $R(T)$ is dense in $L_2(0,1)$ (see Exercise 17.4.3), the functions $y_1(u) \equiv 1$ and $y_2(u) \equiv \sqrt{u}$ are clearly not in $R(T)$.

Thus $R(T)$ is not closed.

This lack of closure of $R(T)$ arises from the simple fact that the inverse operator $T^{-1}y(u) = u^{-1}y(u)$ is not bounded over $y \in L_2(0,1)$ (consider $y = 1$, for example).

On the other hand, it is easy to verify that for any normed spaces $\mathbb{D}$ and $\mathbb{E}$ and any $T \in B(\mathbb{D}, \mathbb{E})$, $N(T)$ is always closed as a direct consequence of the continuity of T.

Observe also that for any $T \in B(\mathbb{B}_1, \mathbb{B}_2)$,

$$\begin{aligned} N(T^*) &= \{b_2^* \in \mathbb{B}_2^* : (T^* b_2^*) b_1 = 0 \text{ for all } b_1 \in \mathbb{B}_1\} \qquad (2) \\ &= \{b_2^* \in \mathbb{B}_2^* : b_2^*(T b_1) = 0 \text{ for all } b_1 \in \mathbb{B}_1\} \\ &= R(T)^\perp, \end{aligned}$$

where $R(T)^\perp$ is an abuse of notation denoteing the linear functionals in $\mathbb{B}_2^*$ that yield zero on $R(T)$.

For Hilbert spaces, the notation is valid because of the isometry between a Hilbert space $\mathbb{H}$ and its dual $\mathbb{H}^*$.

The identity (2) has an interesting extension:

Theorem 17.7 *For two Banach spaces* $\mathbb{B}_1$ *and* $\mathbb{B}_2$ *and for any* $T \in B(\mathbb{B}_1, \mathbb{B}_2)$, $R(T) = \mathbb{B}_2$ *if and only if* $N(T^*) = \{0\}$ *and* $R(T^*)$ *is closed.*

If we specialize (2) to Hilbert spaces $\mathbb{H}_1$ and $\mathbb{H}_2$, we obtain trivially for any $A \in B(\mathbb{H}_1, \mathbb{H}_2)$ that $R(A)^\perp = N(A^*)$.

The following result for Hilbert spaces is also useful:

Theorem 17.9 *For two Hilbert spaces* $\mathbb{H}_1$ *and* $\mathbb{H}_2$ *and any* $A \in B(\mathbb{H}_1, \mathbb{H}_2)$, $R(A)$ *is closed if and only if* $R(A^*)$ *is closed if and only if* $R(A^*A)$ *is closed.*

Moreover, if $R(A)$ *is closed, then* $R(A^*) = R(A^*A)$ *and*

$$A(A^*A)^{-1}A^* : \mathbb{H}_2 \mapsto \mathbb{H}_2$$

is the projection onto $R(A)$.