STAT 597: Functional Data Analysis
Lecture 5 – Chapter 10 – Hilbert spaces

Instructor: Matthew Reimherr

Outline

- Hilbert spaces
- Linear operators

Vector space

A vector space, $\mathcal{V}$, over the real line, $\mathbb{R}$, is a collection of elements which is closed under adition and scalar multiplication:

- For every $x \in \mathcal{V}$ and $y \in \mathcal{V}$ there corresponds a third vector, denoted $x + y$ which is also in $\mathcal{V}$.
- The opperation of addition is commutitive and associative:

$$x + y = y + x \qquad (x + y) + z = x + (y + z).$$

- There exists a unique element denoted $\mathbf{0} \in \mathcal{V}$ which satisfies $x + \mathbf{0} = x$ for all $x \in \mathcal{V}$.
- For each $x \in \mathcal{V}$, there exists an element $-x \in \mathcal{V}$ such that $x + (-x) = \mathbf{0}$.
- For every $a \in \mathbb{R}$ and $v \in \mathcal{V}$ there corresponds a third vector, denoted ax which is also in $\mathcal{V}$.
- The operation of multiplication by a scalar satisfies

$$a(bx) = (ab)x \qquad (a + b)x = ax + bx \qquad a(x + y) = ax + ay.$$

- Scalar multiplication by 1 results in the same vector, $1x = x$.

Vector space

A *linear subspace*, $\mathcal{G}$, of $\mathcal{V}$, is a subset of $\mathcal{V}$ which is also a vector space under the same addition and multiplication operations as $\mathcal{V}$.

We say that vectors $e_1, \ldots, e_d$ in $\mathcal{V}$ are *linearly independent* if

$$a_1 e_1 + \cdots + a_d e_d = 0 \iff a_1 = a_2 = \cdots = a_d = 0.$$

Furthermore, if every element of $\mathcal{V}$ can be expressed as a linear combination of $\{e_i\}$ then we say they are a *basis* for $\mathcal{V}$ and that the *dimension* of $\mathcal{V}$ is d.

Normed vector space

A *normed vector space* is a pair $(\|\cdot\|, \mathcal{V})$ where $\mathcal{V}$ is a vector space and $\|\cdot\|$ is a mapping from $\mathcal{V} \to [0, \infty)$ which satisfies

- $\|ax\| = |a|\|x\|$,
- $\|x + y\| \leq \|x\| + \|y\|$ (Triangle inequality),
- If $\|x\| = 0$ then $x = \mathbf{0}$.

A *Banach space* is a *complete* normed vector space. We say that the normed vector space $(\|\cdot\|, \mathcal{V})$ is complete if every cauchy sequence converges to a point in the space. Recall that a sequence $\{x_n\}$ is called Cauchy if for any $\epsilon > 0$ there exists N such that

$$\|x_n - x_m\| < \epsilon \quad \forall n, m > N.$$

Hilbert spaces

Finally we arrive at *Hilbert spaces*. A Hilbert space, denoted $\mathcal{H}$, is a Banach space whose norm is defined via an inner product: $\| \cdot \|^2 = \langle \cdot, \cdot \rangle$. An inner product is a mapping from the cartesian product $\mathcal{H} \times \mathcal{H} \to \mathbb{R}$ which satifies:

- $\langle x, y \rangle = \langle y, x \rangle$ (for real Hilbert spaces),
- $\langle a_1 x_1 + a_2 x_2, y \rangle = a_1 \langle x_1, y \rangle + a_2 \langle x_2, y \rangle$,
- $\langle x, x \rangle \geq 0$ with equality only when $x = \mathbf{0}$.

Hilbert spaces might seem only slightly different than Banach spaces, but the mathematics becomes much simpler. Most properties you can think of from scalar/vector settings generalize well to Hilbert spaces, not necessarily Banach spaces.

Example - $\mathbb{R}^d$

Let $\mathbb{R}^d$ be the set of all d dimensional vectors with coordinates in $\mathbb{R}$ and $d < \infty$. The Euclidian norm is given by

$$|x|^2 = \sum_{i=1}^{d} x_i^2.$$

Under this norm, $\mathbb{R}^d$ is a Hilbert space. More generally, one, for $p \geq 1$, can define the ℓ_p norm as

$$|x|_p^p = \sum_{i=1}^{d} x_i^p \qquad |x|_\infty = \max_{1 \leq i \leq p} |x_i|.$$

Under any of these norms, $\mathbb{R}^d$ is only a Banach space unless $p = 2$.

Example - $\mathcal{C}[0,1]$

The space of continuous functions over $[0,1]$ is often denoted $\mathcal{C}[0,1]$. Equipped with the *sup-norm* this space is a Banach space:

$$\|x\| = \sup_{0 \leq t \leq 1} |x(t)|.$$

However the sup-norm is not an inner product norm (why?), and thus this is not a Hilbert space.

Example - $L^2[0,1]$

This space is the "bread and butter" of many FDA methods. The space $L^2[0,1]$ is the space of real valued functions over $[0,1]$ which are square integrable:

$$\|x\|^2 = \int_0^1 x(t)^2 \, dt.$$

Showing that it is an inner product space (i.e. a Hilbert space that is not necessarily complete) is straightfoward, showing that the space is complete takes more work.

Other examples

- Reproducing kernel Hilbert spaces,
- Sobolev spaces,
- Tensor product and Cartesian product spaces.

Projections

A major tool in FDA is functional principal component analysis, which involves projecting the data onto lower dimensional subspace. An advantage of using Hilbert spaces is that the inner product actually defines the geometry of the space. For example, we say that two elements are *perpendicular* if $\langle x, y \rangle = 0$. We will review some facts about projections.

Uniqueness of Projections

Theorem

Let $\mathcal{G}$ be a closed subspace of $\mathcal{H}$. For any $x \in \mathcal{H}$ define

$$\delta_x = \inf_{z \in \mathcal{G}} \|x - z\|.$$

Then there exists a unique point $y \in \mathcal{G}$ which achives $\|y - x\| = \delta_x$.

The point y is called the projection of x onto $\mathcal{G}$, and is denote $P_{\mathcal{G}}(x)$ or just $P(x)$ for short.

Properties

The orthogonal complement of $\mathcal{G}$, denoted $\mathcal{G}^\perp$ is defined as

$$\mathcal{G}^\perp = \{x \in \mathcal{H} : \langle x, y \rangle = 0 \text{ for some } y \in \mathcal{G}\}.$$

Note that $\mathcal{G}^\perp$ is also a closed linear subspace (why?). Denote $Q(\cdot)$ as the projection onto $\mathcal{G}^\perp$.

- Every $x \in \mathcal{H}$ can be uniquely decomposed as $x = P(x) + Q(x)$.
- $P(x)$ and $Q(x)$ are nearest points to x in $\mathcal{G}$ and $\mathcal{G}^\perp$.
- The mappings $P(\cdot)$ and $Q(\cdot)$ are linear, i.e.

$$P(ax + by) = aP(x) + bP(y).$$

- The norm of x satisfies $\|x\|^2 = \|P(x)\|^2 + \|Q(x)\|^2$.

Cauchy-Schwarz inequality

A very useful result for Hilbert spaces is the Cauchy-Schwarz inequality:

$$|\langle x, y\rangle| \leq \|x\|\|y\|.$$

It is a fairy simple proof. Let $e = \|x\|^{-1}x$, then

$$\begin{aligned}\|y\|^2 &= \|y - \langle y, e\rangle e\|^2 + \|\langle y, e\rangle e\|^2 \\ &\geq \langle y, e\rangle^2 = \frac{\langle y, x\rangle^2}{\|x\|^2}.\end{aligned}$$

So, as desired

$$\|x\|^2\|y\|^2 \geq \langle x, y\rangle^2.$$

(Note $P_x(y) = \langle e, x\rangle e$ and $Q(x) = y - P(x)$)

Riesz Representation Theorem

Theorem

If $L : \mathcal{H} \to \mathbb{R}$ is continuous and linear then there exists $y \in \mathcal{H}$ such that

$$L(x) = \langle x, y \rangle \qquad \forall x \in \mathcal{H}.$$

Cleary the function $\langle \cdot, y \rangle$ is continuous and linear, but the Riesz Representation Theorem states that every continuous linear functional is actually of this form. (Note: we will use the term "functional" to denote a mapping to the real line).

For any Hilbert (or Banach space), the set of continuous linear functionals is called the *Dual Space* and usually denoted as $\mathcal{H}^*$. The Riesz Representation Theorem states that, in some sense $\mathcal{H} = \mathcal{H}^*$.

Basis Expansions

Suppose that $e_1, \ldots, e_d$ is an orthonormal basis for a subspace $\mathcal{G} \subset \mathcal{H}$. This means that for any $x \in \mathcal{H}$ we have $P(x) \in \mathcal{G}$ and so

$$P(x) = a_1 e_1 + \cdots + a_d e_d \quad \text{and} \quad \langle e_i, e_j \rangle = 1_{i=j}.$$

Using our discussed properties we have

$$\begin{aligned}
\langle x, e_i \rangle &= \langle P(x) + Q(x), e_i \rangle \\
&= \langle P(x), e_i \rangle + \langle Q(x), e_i \rangle \\
&= \langle a_1 e_1 + \cdots + a_d e_d, e_i \rangle = a_i.
\end{aligned}$$

This means that $a_i = \langle x, e_i \rangle$ and so the projection, $P(\cdot)$, can be written as

$$P(x) = \sum_{i=1}^{d} \langle x, e_i \rangle e_i.$$

Separability

We say that the space $\mathcal{H}$ is separable if it contains a countable basis (there is a more general defintion for other spaces). This means that any $x \in \mathcal{H}$ can be expressed as

$$x = \sum_{i=1}^{\infty} a_i e_i.$$

Theorem (Parceval's Theorem)

Let $\{e_i\}$ be an orthnormal basis of a real separable Hilbert space, $\mathcal{H}$. Then for any $x \in \mathcal{H}$ we have

$$x = \sum_{i=1}^{\infty} \langle x, e_i \rangle e_i \quad \text{and} \quad \|x\|^2 = \sum_{i=1}^{\infty} \langle x, e_i \rangle^2.$$

Nonseparable Banach spaces are uncommon, but not so difficult to construct. Nonseparable Hilbert spaces, however, are fairly exotic.

Linear operators

Recall that an operator $L : \mathcal{H} \to \mathcal{H}$ is called linear if $L(ax + by) = aL(x) + bL(y)$. We say the linear operator is *bounded* if

$$\|L\|_{\mathcal{L}} = \sup_{\|x\| \le 1} \|L(x)\| < \infty.$$

We denote the space of all bounded operators as $\mathcal{L}$. Under the norm $\| \cdot \|_{\mathcal{L}}$, this space is a Banach space.

Note that a classic problem from real analysis is to show that a linear operator is bounded if and only if it is continuous (if and only if it is continuous at a point).

Operator inequality

By definition of $\|L\|_{\mathcal{L}}$ one has the the fairly trivial *operator inequality*

$$\|L(x)\| \leq \|L\|_{\mathcal{L}}\|x\|.$$

Simple, but like Cauchy-Schwarz, it can be very useful.

Hilbert-Schmidt operators

A bounded linear operator, L, is called Hilbert-Schmidt if

$$\|L\|_{\mathcal{S}}^2 = \sum_{i=1}^{\infty} \|L(e_i)\|^2 < \infty.$$

where e_i is some orthnormal basis. The above quantity is invariant with respect to the choice of basis. We denote the space of Hilbert-Schmidt operators as $\mathcal{S}$, equipped with the norm $\|\cdot\|_{\mathcal{S}}$ it is a Hilbert space.

Example - Identity operator

The identity operator, L, is defined as $L(x) = x$. This operator is bounded since

$$\|L\|_{\mathcal{L}} = \sup_{\|x\|\leq 1} \|L(x)\| = \sup_{\|x\|\leq 1} \|x\| = 1.$$

However, it is not Hilbert-Schidt

$$\|L\|_{\mathcal{S}}^2 = \sum_{i=1}^{\infty} \|L(e_i)\|^2 = \sum_{i=1}^{\infty} \|e_i\|^2 = \sum_{i=1}^{\infty} 1 = \infty.$$

Example - Integral operator

Suppose that $\mathcal{H} = L^2[0,1]$. Consider the bivariate function $\psi(t,s)$ which satisfies

$$\int\int \psi(t,s)^2 \; dtds < \infty.$$

This function can be used to define an *integral operator*

$$\Psi(x)(t) = \int \psi(t,s)x(s) \; ds.$$

This operator is bounded and Hilbert-Schmidt.

Example - Integral operator

Let $\|x\| \leq 1$ then by the Cauchy-Scwartz inequality

$$\begin{aligned}
\|\Psi(x)\|^2 &= \int \left(\int \psi(t,s)x(s)\ ds \right)^2 dt \\
&\leq \int \left(\int \psi(t,s)^2\ ds \right) \left(\int x(s)^2\ ds \right) dt \\
&\leq \int \int \psi(t,s)^2\ dsdt.
\end{aligned}$$

So $\Psi \in \mathcal{L}$ (can you guess what $\|\Psi\|_{\mathcal{L}}$ is?)

Example - Integral operator

Let $\psi_t(s) = \psi(t, s)$, then ψ_t is an element of $L^2[0, 1]$. We have

$$
\begin{aligned}
\|\Psi\|_{\mathcal{S}}^2 &= \sum \|\Psi(e_i)\|^2 \\
&= \sum \int \left(\int \psi(t, s) e_i(s) \; ds \right)^2 \; dt \\
&= \int \sum \langle \psi_t, e_i \rangle^2 \; dt \\
&= \int \|\psi_t\|^2 \; dt = \int \int \psi(t, s)^2 \; dtds.
\end{aligned}
$$

So $\Psi \in \mathcal{S}$.

Bounded vs Hilbert-Schmidt

It is important to remember that the Hilbert-Schmidt property is stronger than being bounded, i.e. $\mathcal{S} \subset \mathcal{L}$ and we have

$$\|\Psi\|_{\mathcal{L}} \leq \|\Psi\|_{\mathcal{S}} \qquad \text{for all } \Psi \in \mathcal{S}.$$

The proof involves a basis expansion:

$$\Psi(x) = \sum \langle x, e_i \rangle \Psi(e_i)$$

followed by the triangle inequality and Cauchy-Schwarz:

$$\begin{aligned} \|\Psi(x)\| &\leq \sum |\langle x, e_i \rangle| \|\Psi(e_i)\| \\ &\leq \left(\sum |\langle x, e_i \rangle|^2 \right)^{1/2} \left(\sum \|\Psi(e_i)\|^2 \right)^{1/2} = \|x\| \|\Psi\|_{\mathcal{S}}. \end{aligned}$$