Functional Analysis

Daniele Zago

November 25, 2021

Contents

Lecture 5: Hilbert spaces

2021-11-23

We now discuss Hilbert spaces, which are spaces where it is possible to define the notions of length and orthogonality. With these definitions, we are able to work with the elements of the space geometrically, as if they were vectors in Euclidean space.

Def. (Closed subspace)
X Banach space and $V \subseteq X$ is a ***closed subspace of X*** if

1. $V \subseteq X$ and V is a vectorial space.
2. If $f_n \in V$ for al n such that $f_n \to f$ in X, then $f \in V$.

V contains all the limit points of its converging sequences.

Example (M^2)
Let $V = \{X \in M^2 : \mathbb{E}[X] = 0\}$.

1. V is a subspace of M^2, indeed if $X_1, X_2 \in V$ then also $\alpha X_1 + \beta X_2 \in V$.
2. Check that every converging sequence in V we have a limit in V itself.
   $X_n \xrightarrow{L^2} X$ is such that $\mathbb{E}[X_n] = 0$, since $X_n \in V$ and $\mathbb{E}[(X_n - X)^2] \to 0$ since $V \subseteq X$. We want to deduce that $\mathbb{E}[X] = 0$.

$$\mathbb{E}[|X_n - X|] \overset{??}{\leq} \sqrt{\mathbb{E}[(X_n - X)^2]},$$

   therefore $\mathbb{E}[|X_n - X|] \to 0$, but now apply the expected value operator to all members of the following inequality to get convergence of X_n to X in V,

$$-|X_n - X| \leq X_n - X \leq |X_n - X|$$

Def. (Hilbert space)
Let X be a Banach space, then X is a ***Hilbert space*** if on X it is also defined a *scalar product* $\langle \cdot, \cdot \rangle : X \times X \longrightarrow \mathbb{R}$.

Properties of a scalar product

1. *Positive*: $\langle x, y \rangle \geq 0$ for all $x, y \in X$.
2. *Symmetric*: $\langle x, y \rangle = \langle y, x \rangle$ for all $x, y \in E$.
3. *Bilinear*: $\langle ax, y \rangle = a\langle x, y \rangle$ for all $x, y \in E$ and $a \in \mathbb{R}$..

Example (Scalar product on M^2)
If we consider M^2, then $\|X\|_2 = \sqrt{\mathbb{E}[|X|^2]}$ and $X \cdot Y \in L^1$ by Hölder's inequality. Is the norm and the scalar product can be defined as

$$\langle X, Y \rangle = \mathbb{E}[X \cdot Y] \in \mathbb{R}.$$

M^k for $k \neq 2$ does not guarantee that $X \in M^k$ and $Y \in M^k$ implies $XY \in M^1$.

Example (L^2)
Analogously, for L^2 we can define for $f, g \in L^2$ since $fg \in L^1$ the scalar product defined by

$$\langle f, g \rangle = \int_{\mathbb{R}} f(x)g(x)\,\mathrm{d}x.$$

Prop. 1 (Scalar product associated with a norm)
*A norm $\|\cdot\|$ on a Banach space is associated with a scalar product $\langle \cdot, \cdot \rangle$, i.e. $\langle x, x \rangle = \|x\|^2$ $\iff$ it satisfies the **parallelogram identity**,*

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2, \quad \forall x, y \in X.$$

Thm. 1 (Cauchy-Schwartz inequality)
Let $x, y \in X$, then $|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$.

With Hilbert spaces we can now define a notion of *orthogonality between vectors* $x, y \in H$.

Def. (Orthogonality)
$x, y \in H$ are ***orthogonal*** if $\langle x, y \rangle = 0$.

Example ($L^2(-1, 1)$)
Consider $L^2(-1,1) = \{f : \mathbb{R} \to \mathbb{R} \text{ such that } \int_{-1}^{1} |f(x)|^2\,\mathrm{d}x < \infty\}$.
We can take $f(x) = \sin x$ and $g(x) = 1$, and

$$\langle f, g \rangle = \int_{-1}^{1} \sin x\,\mathrm{d}x = 0.$$

Def. (Orthogonal complement)
Let $S \subseteq H$, we define the ***orthogonal complement*** $S^\perp$ of S as the set of orthogonal elements to S,

$$S^\perp = \{h \in H : \langle h, s \rangle = 0,\ \forall s \in S\}.$$

Subspace The orthogonal complement $S^\perp$ is a subspace of S since if $h_1, h_2 \in S^\perp \implies h_1 + h_2 \in S^\perp$ by bilinearity of the scalar product.

Closedness Moreover, $S^\perp$ is also closed since by the Cauchy-Schwartz inequality (thm. 1) the scalar product is continuous w.r. to convergence of elements.

Double complement We have that $(S^\perp)^\perp = \{h \in H : \langle h, v \rangle,\ \forall v \in S^\perp\}$, therefore $S \subseteq (S^\perp)^\perp$. However, $(S^\perp)^\perp$ is a closed subspace of H, and if S is itself a closed subspace of H then $S = (S^\perp)^\perp$. In general, $(S^\perp)^\perp$ is the smallest closed subspace of H which contains S.

Thm. 2 (Orthogonal projection)
Let H be a Hilbert space and let $V \subseteq H$ be a closed subspace of H. Then, $H = V + V^\perp$, in the sense that for each $h \in H$ there exists an unique $v \in V$ and an unique $w \in V^\perp$ such that

$$h = v + w.$$

Orthogonal projection This theorem states that v is the *orthogonal projection* of h onto V, which is the element in V which has ***minimal distance*** from h,

$$v = \operatorname*{argmin}_{z \in V} \|h - z\|.$$

Proof.
We want to solve the minimization problem assuming $h \notin V$, otherwise $v = h$,

$$\min_{z \in V} \|z - h\|.$$

We define

$$\delta := \inf_{z \in V} \|h - z\|,$$

and we take a sequence $v_n \in V$ such that $\delta \leq \|h - v_n\| \leq \delta + \frac{1}{n}$. Since we are in a Hilbert space, using the parallelogram identity we have that

$$\|v_n - v_m\| \xrightarrow{n,m \to \infty} 0,$$

therefore $v_n \xrightarrow{n \to \infty} v \in V$ since V is closed and

$$\delta \leq \|h - v\| \leq \delta.$$

$\square$

Def. (Orthonogonal projection operator)
For all $V \subseteq H$ closed subset of H we define the ***orthogonal projection operator*** P_V as the operator that projects $h = v + w$ onto the space V,

$$h \longmapsto P_V(h) = v.$$

Lecture 6: Hilbert space (cont.)

2021-11-24

We are interested in finding a solution to the problem of finding the projection $P_V(x)$ of an element x of the Hilbert space X, where V is a closed subset of X.

6.1 Orthogonal projections

Example (Conditional distribution)
Consider for instance the problem of finding the minimizer over

$$\mathbb{E}[(X - \mathbb{E}[X|V])^2] = \min_{z \in V} \mathbb{E}[(X - Z)^2],$$

therefore $\mathbb{E}[X|V]$ is the orthogonal projection of X over the subspace.

Example (Linear least squares estimation)
Given $X, Y \in M^2$ we want to find $a, b \in \mathbb{R}$ such that

$$\operatorname*{argmin}_{a,b \in \mathbb{R}} \mathbb{E}[(Y - a - bX)^2].$$

We can rewrite this problem as an orthogonal projection over a Hilbert space M_2 given the closed subset $V = \{Y \in M_2 : Y = a + bX, \text{ for some } a, b \in \mathbb{R}\}$. Given $Y \in M_2$ we define the orthogonal projection of Y

$$P_V(Y) = \mathbb{E}[Y|V],$$

as the element in V with minimal distance from Y,

$$P_V(Y) = a + bX, \quad \text{for some } a, b \in \mathbb{R},$$

such that

$$\mathbb{E}[(Y - P_V(Y))^2] = \min_{Z \in V} \mathbb{E}[(Y - Z)^2] = \min_{\alpha, \beta \in \mathbb{R}} \mathbb{E}[(Y - \alpha - \beta X)^2].$$

Therefore, the problem is how to compute the orthogonal projection on V.

We want a method to compute the orthogonal projection of elements in a Hilbert space when given a subspace V. In general, we can use two approaches in order to perform the required operation:

1. Write down the objective function and minimize it, as in the examples above.
2. Use the concept of a *basis* as in standard linear algebra.

Def. (Orthonormal set)
Let H be a Hilbert space, then a set $\{e_i, i \in \mathbb{N}\}$ is called an ***orthonormal set*** if

$$\langle e_i, e_j \rangle = \delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \quad \text{for any } i, j \in I.$$

Example ($\mathbb{R}^n$)
In $\mathbb{R}^n$ an orthonormal set is $\{(1,0,\dots,0),(0,1,\dots,0),\dots,(0,0,\dots,1)\}$.

Thm. 3 (Orthonormal basis)
Let H be a Hilbert space and $\{e_i, i \in I\}$ ba an orthonormal set of elements of H. Then, the following conditions are equivalent

1. *If for some $h \in H$ we have $\langle h, e_i \rangle = 0$ for all $i \in I$, then $h = 0$.*
2. *For every $h \in H$, then $h = \sum_{i \in I} \langle h, e_i \rangle e_i$.*

Orthonormal basis If *1.* or *2.* hold, then we say that $\{e_i, i \in I\}$ is an ***orthonormal basis*** of H.

Approximation We have an approximation given by

$$\lim_{N\to\infty} \|h - \sum_{i=1}^{N} \langle h, e_i \rangle e_i\|,$$

and we have that the norm of h is equal to

$$\begin{aligned} \|h\|^2 &= \langle h, h \rangle \\ &= \langle \sum_i \langle h, e_i \rangle e_i, \sum_j \langle h, e_j \rangle e_j \rangle \\ &= \sum_{i,j} \langle h, e_i \rangle \langle h, e_j \rangle \langle e_i, e_j \rangle \\ &= \sum_{i=j} \langle h, e_i \rangle \langle h, e_j \rangle. \end{aligned}$$

Thm. 4
Every Hilbert space admits an orthonormal basis. Moreover, if the orthonormal basis is countable ($|I| = |N|$), then H is called a ***separable*** *Hilbert space.*

Remark The sum of an uncountable set can be interpreted first by saying that $I = \mathbb{R}$ means

$$h = \sum_{i \in I} \langle h, e_i \rangle e_i \iff \langle h, e_i \rangle \neq 0 \text{ only for a countable } J \subseteq \mathbb{N} \text{ and } h = \sum_{j \in J} \langle h, e_i \rangle e_i.$$

Let $V \subseteq H$ be a subspace of H, then V is itself a Hilbert space with the same scalar product of H and it is closed. Then, we can find an orthonormal basis of V given by the orthonormal set of vectors $\{v_i : i \in I\}$ and for every $h \in H$ we have the orthogonal projection

$$P_V(h) = \sum_i \langle h, v_i \rangle v_i.$$

Example (Linear least squares)
back to our example, $V = \{\alpha + \beta X, \alpha, \beta \in \mathbb{R}\}$. Then, a basis of V is simply $\{1, X\}$, which however is not orthonormal. Firstly, we want them to be orthogonal and

$$\langle 1, X \rangle = \mathbb{E}[X] \neq 0 \text{ in general,}$$

therefore we replace X by $X - \mathbb{E}[X]$ and $\{1, X - \mathbb{E}[X]\}$ is still a basis of V. Finally, we normalize each vector by their norm

$$\|1\|_2 = \sqrt{\mathbb{E}[(1)^2]} = 1$$
$$\|X - \mathbb{E}[X]\|_2 = \sqrt{\mathbb{E}[(X - \mathbb{E}[X])^2]} = \sqrt{\mathbb{V}[X]}.$$

Therefore we finally obtain the orthonormal basis

$$\left\{1, \frac{X - \mathbb{E}[X]}{\sqrt{\|\mathbb{V}[X]\|}}\right\},$$

and the orthogonal projection $P_V(Y)$ of Y onto V is simply

$$\begin{aligned} P_V(Y) &= \sum_{i=1}^{2} \langle Y, v_i \rangle v_i \\ &= \langle Y, 1 \rangle \cdot 1 + \langle Y, \frac{Y - \mathbb{E}[X]}{\mathbb{V}[X]} \rangle \cdot \frac{X - \mathbb{E}[X]}{\mathbb{V}[X]} \\ &= \mathbb{E}[Y] + \frac{\mathbb{E}[XY] - \mathbb{E}[X]}{\sqrt{\mathbb{V}[X]}} \cdot \frac{X - \mathbb{E}[X]}{\sqrt{\mathbb{V}[X]}} \\ &= \mathbb{E}[Y] + \frac{\mathrm{Cov}(X, Y)}{\mathbb{V}[X]} (X - \mathbb{E}[X]). \end{aligned}$$

We now generalize to the infinite-dimensional case the properties of linear operators seen in linear algebra.

$$\begin{aligned} T : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ x &\longmapsto Ax \end{aligned} \quad \xrightarrow{n \to \infty} \quad \begin{aligned} T : H &\longrightarrow H \\ f &\longmapsto Tf \end{aligned}$$

We define the adjoint of T as the operator T^* such that

$$\langle Tx, y \rangle = \langle x, T^* y \rangle, \quad \text{for all } x, y \in H.$$

Indeed this generalizes the transpose, since $(Ax)^\top y = x^\top A^\top y$.
We consider the operator

$$\begin{aligned} T : L^2(0,1) &\longrightarrow L^2(0,1) \\ f &\longmapsto Tf(x) = \int_0^x f(t)\, \mathrm{d}t. \end{aligned}$$

Firstly we have to prove that if $f \in L^2(0,1)$ then $Tf \in L^2(0,1)$ since

Therefore, $Tf \in L^2$ and t is linear (integral is linear) and bounded since by the computation above we have

$$\|Tf\|_2^2 \leq \frac{1}{2}\|f\|_2^2.$$

In order to compute the adjoint we have to find a T^* such that

$$\langle Tf, g\rangle = \int_0^1 Tf(x)g(x)\,\mathrm{d}x = \int_0^1 f(x)T^*g(x)\,\mathrm{d}x = \langle f, T^*g\rangle.$$

The two terms have to be equal, i.e.

$$\begin{aligned}
\int_0^1 f(x)T^*g(x) &= \int_0^1 Tf(x)g(x)\,\mathrm{d}x \\
&= \int_0^1 \int_0^x f(t)\,\mathrm{d}t g(x)\,\mathrm{d}x && \text{(def.)} \\
&= \int_0^1 \int_t^1 f(t)g(x)\,\mathrm{d}x\,\mathrm{d}t && (0 \leq t \leq x \leq 1) \\
&= \int_0^1 f(t)\left(\int_t^1 g(x)\,\mathrm{d}x\right)\,\mathrm{d}t
\end{aligned}$$

Therefore, we conclude that

$$T^*g(t) = \int_t^1 g(x)\,\mathrm{d}x.$$

Def. (Self adjoint operator)
If $T^* = T$ we say that T is a ***self-adjoint*** operator.

Def. (Eigenfunction)
Given an operator $T : H \to H$, we say that $\lambda \in \mathbb{R}$ is an ***eigenvalue*** of T if there exists an $h \in H$ such that

$$Th = \lambda h,$$

and h is called ***eigenfunction*** (or ***eigenvector***).

Eigenvalues of T are contained in the so-called ***spectrum*** of T, that is

$$\{\lambda \in \mathbb{R}\text{such that } (T - \lambda I) \text{ is not invertible}\},$$

where

$$\begin{aligned}
T - \lambda I : H &\longrightarrow H \\
h &\longmapsto Th - \lambda h.
\end{aligned}$$

If λ is an eigenvalue of T then there exists $h \in H$ such that $(T - \lambda I)(h) = 0$. In the finite-dimensional case, the set of eigenvalues coincides with the spectrum, whereas in the infinite-dimensional case they form the ***point spectrum*** of T.

Def. (Compact linear operator)
Operators for which the spectrum is given by the set of eigenvalues are called ***compact linear operators***.

Prop. 2 (Compact operators compress the space)
An operator $T : H \longrightarrow H$ is compact operator if and only if $h_n \in H$ is a bounded sequence (not necessarily converging), then $Th_n \to Th$.

A compact operator in L^2 is typically of the form

$$f \longrightarrow Tf(x) = \int_A K(x,y)f(y)\,\mathrm{d}y,$$

and in the example above this was exactly

$$f \longrightarrow Tf(x) = \int_0^1 \mathbb{1}_{[0,x]}(y)f(y)\,\mathrm{d}y.$$

Symmetric compact operator are diagonalizable just like symmetric matrices in the finite-dimensional case.

Prop. 3 (Existence of an orthonormal basis)
If $T : H \longrightarrow H$ a is compact and symmetric operator on H separable Hilbert space, then there exists an orthonormal basis $\{e_i, i \in I\}$ of H made of eigenfunctions of T, i.e.

$$Te_i = \lambda e_i, \quad \textit{for all } i \in I.$$

Moreover, $\{\lambda_i, i \in I\}$ can be either a finite set (i.e. some of them are equal) or an infinite set $\{\lambda_i\}$ and satisfies

$$\lim_{i\to\infty} \lambda_i = 0.$$

Among compact operators we have a particular subclass which are called Hilbert-Schmidt (or trace-class) operators which satisfy

$$\{\lambda_i\} \lim_{i\to\infty} \lambda_i \quad \text{and} \quad \sum_{i=1}^{\infty} |\lambda_i|^2 < \infty.$$

Example (Continued)
Let $Tf(x) = \int_0^x f(t)\,\mathrm{d}t$, then T is linear bounded from before. Moreover, it can be proven that T is compact but not symmetric, therefore we want to compute its eigenvalues if they exist: we want to find $\lambda \in \mathbb{R}$ such that

$$Tf(x) = \lambda f(x) \quad \text{for all } x,$$

therefore

$$\int_0^x f(t)\,\mathrm{d}t = \lambda f(x) \iff f \text{ continuous and differentiable,}$$

therefore if we differentiate both sides we find

$$\frac{\partial}{\partial x}\int_0^x f(t)\,\mathrm{d}t = \frac{\partial}{\partial x}\lambda f(x) \iff \begin{cases} f(x) = \lambda f'(x) \\ 0 = \lambda f(0) \end{cases}$$

therefore $f(x) = e^{x/\lambda} \implies f'(x) = \frac{1}{\lambda}e^{x/\lambda}$. However, there $\lambda e^0 = \lambda$ but since $\lambda = 0$ we have $f = 0$ is not an eigenvalue. Therefore, T has no eigenvalues.