Hilbert Space, Fourier and Learning Machines

Artur Sahakyan

April 6, 2025

1 Introduction

The word *space* has multiple meanings and applications. Quite commonly, it refers to the three-dimensional world we live in. For some, it might first evoke the cosmos, a workspace, or even a designated area within a chatbot. In mathematics, however, the concept of space takes on a more abstract form. A space can be a set of possibilities, a collection of events, a group of elements that follow certain rules, or a set of functions with specific properties.

A *Hilbert space*, for instance, has its own defining characteristics. It is an abstract space, and some of its elements may also have real-world representations. One important aspect of Hilbert space is that it can be infinite-dimensional, meaning the set of linearly independent vectors it can contain is infinite.

To better understand what Hilbert space is, let's explore a few foundational concepts along with their definitions.

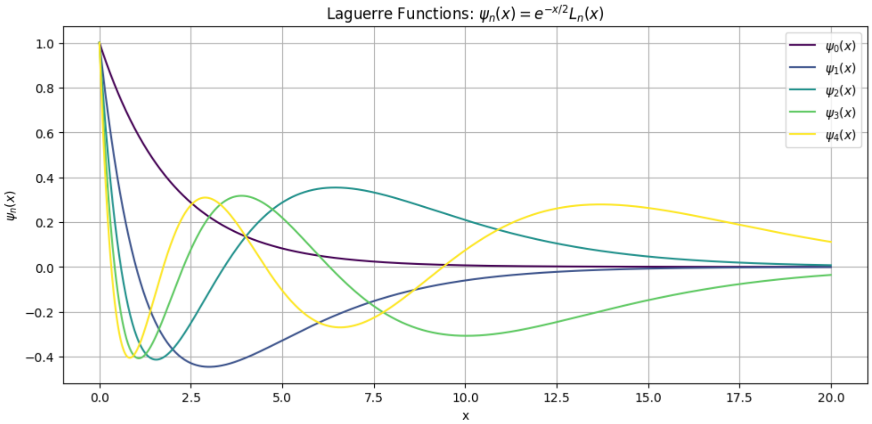


Figure 1: Laguerre functions which are in a Hilbert space

2 Metric Space

Formally, a metric space is an ordered pair (M,d), M is a set with d being a metric on it.

Let M be a set. A function

$$d : M \times M \to \mathbb{R}$$

is called a **metric** (or distance function) on M if for all $x, y, z \in M$, the following properties hold:

1. **Nonnegativity:**
$$d(x, y) \geq 0$$

2. **Nondegeneracy:**
$$d(x, y) = 0 \iff x = y$$

3. **Symmetry:**
$$d(x, y) = d(y, x)$$

4. **Triangle inequality:**
$$d(x, z) \leq d(x, y) + d(y, z)$$

If a function d satisfies these conditions, then the pair (M, d) is called a *metric space*. We can define distance function in many ways, so long as the properties of the metric space are preserved.

For instance, this can be a distance function:

$$d(x, y) = \begin{cases} 0 & \text{if } x = y, \\ 1 & \text{if } x \neq y. \end{cases}$$

3 Normed Space

A normed space is a vector space with a metric defined by the norm.

Let V be a linear space over the field $\mathbb{K}$, where $\mathbb{K}$ is either $\mathbb{R}$ (a real number) or $\mathbb{C}$ (a complex number). A norm on V is a function

$$\|\cdot\| : V \to \mathbb{R}$$

that satisfies the following properties for all $v, w \in V$ and all scalars $\lambda \in \mathbb{K}$:

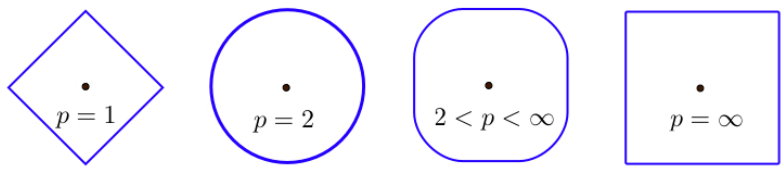
1. **Nonnegativity:**
$$\|v\| \geq 0$$
2. **Nondegeneracy:**
$$\|v\| = 0 \iff v = 0$$
3. **Multiplicality (absolute scalability):**
$$\|\lambda v\| = |\lambda| \cdot \|v\|$$
4. **Triangle inequality:**
$$\|v + w\| \leq \|v\| + \|w\|$$

If $\|\cdot\|$ satisfies all these conditions, then the pair $(V, \|\cdot\|)$ is called a *normed space*.

To understand norms better, let's take a look at the formula for p-norms, with $p \geq 1$.

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

Norms can induce metric, let's observe unit spheres for different metrics in Figure 2.



The unit sphere for different metrics: $\|x\|_{l_p} = 1$ in $\mathbb{R}^2$.

Figure 2: Unit spheres for different metrics, that can be induced by norms

4 Inner Product Space

Inner product spaces are special type of normed spaces. By its definition, it is a vector space X that has inner product defined on it.

Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$. An **inner product** on V is a function

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$$

(where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) satisfying the following properties for all $u, v, w \in V$ and $\lambda \in \mathbb{K}$:

1. **Nonnegativity:** $\langle v, v \rangle \geq 0$
2. **Nondegeneracy:** $\langle v, v \rangle = 0 \iff v = 0$
3. **Multiplicativity:**
$$\langle \lambda v, w \rangle = \lambda \langle v, w \rangle$$
4. **Symmetry (or Hermitian symmetry over $\mathbb{C}$):**
$$\begin{aligned} &\langle v, w \rangle = \langle w, v \rangle \\ \text{or} \quad &\langle v, w \rangle = \overline{\langle w, v \rangle} \end{aligned}$$

5. **Distributivity:**

$$\langle v, w + u \rangle = \langle v, w \rangle + \langle v, u \rangle$$

The norm induced by an inner product is defined as:

$$\|v\| := \sqrt{\langle v, v \rangle}$$

5 Completeness

The property of space being complete is important to define, as it is key in the upcoming definition for Hilbert Space. At the same time, it is crucial to distinguish between the completeness of a metric space, and the completeness of a function space. This will also help us define the big L^2 space.

1. Completeness of a Metric Space

A metric space (M, d) is said to be **complete** if every Cauchy sequence in M converges to a point in M. That is,

$$\forall (x_n) \subset M, \quad (\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall m, n \geq N,\ d(x_m, x_n) < \varepsilon) \Rightarrow \exists x \in M,\ \lim_{n \to \infty} x_n = x$$

Cauchy sequences, put in simple terms, after some point N, have elements getting closer and closer to each other, until eventually they converge to some point (Figure 3). If that point is inside the metric space M, then we say M is complete.

2. Completeness of a System of Functions

A set of functions $\{f_k\} \subset L^2(X)$ is said to be **complete** (or dense in L^2) if:

$$\forall f \in L^2(X),\ \forall \varepsilon > 0,\ \exists c_1, c_2, \ldots, c_n \in \mathbb{R} \text{ such that } \left\| f - \sum_{k=1}^{n} c_k f_k \right\|_{L^2} < \varepsilon$$

Equivalently, $\{f_k\}$ is complete if every function in $L^2(X)$ can be approximated arbitrarily well (in L^2-norm) by finite linear combinations of f_k.

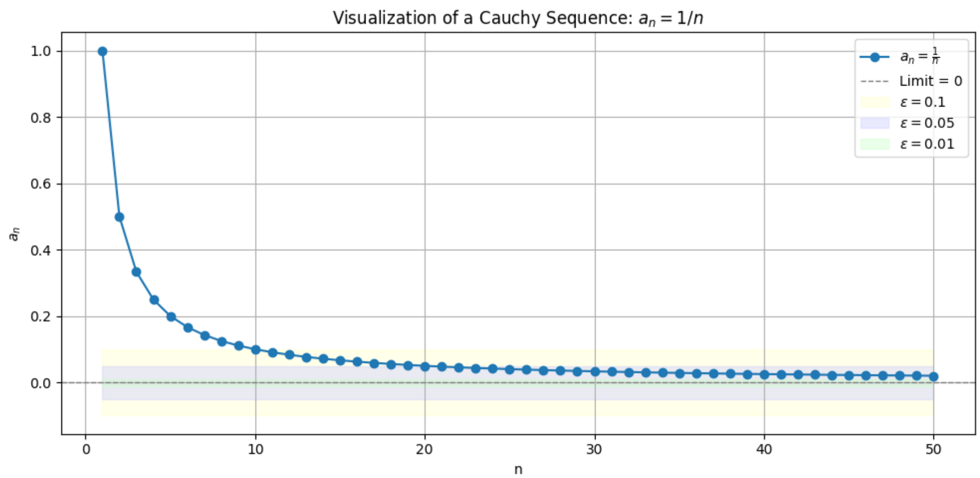


Figure 3: Cauchy sequence example

Important Distinction

- Completeness of a space: Every Cauchy sequence converges in space.
- Completeness of a system: The span is dense in the space.

6 Hilbert Space

Finally, let us give a formal definition for Hilbert Space.

Hilbert space is a complete inner product space (complete in the metric defined by the inner product).

As a side note, if a normed space is complete, it is called a **Banach space**.

7 Euclidean Space

The Euclidean space is finite-dimensional and is a subset of Hilbert space.

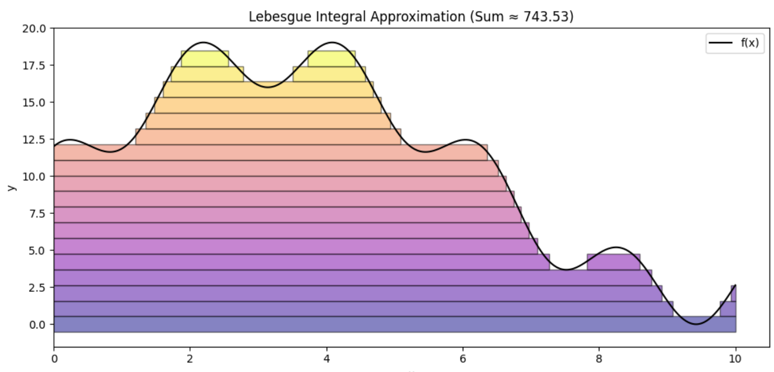
Every Euclidean space is a Hilbert space, but not every Hilbert space is an Euclidean space.

8 Lebesgue Integral

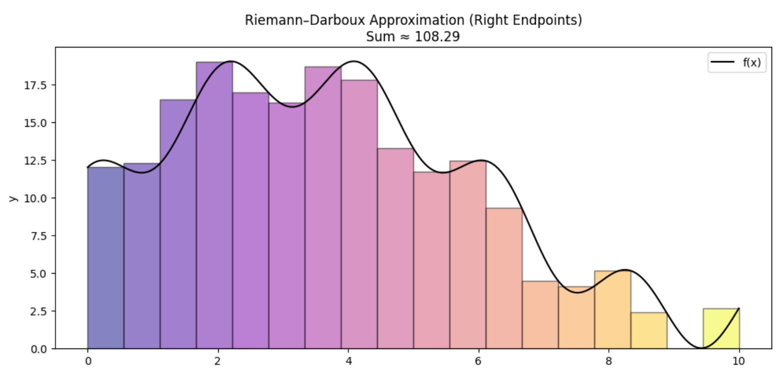
To provide another example of space in Hilbert spaces, let us introduce the Lebesgue integral. In short, it is generalization of the Riemann integral, and instead of taking vertical slices, we consider horizontal slices of function and compute the integral with them (see Figure 4). It is also a more versatile tool, handles discontinuities better, and takes into account the measurability of a function. That's why it is defined as

$$\int_E f \, dm$$

over the region E with respect to measure m.



(a) Lebesgue Integral



(b) Riemann Integral

Figure 4: Comparison of the Lebesgue and Riemann Integrals

9 L^2 space and l^2 space

Let $(X, \mathcal{M}, \mu)$ be a measure space (e.g., $X = [a, b] \subset \mathbb{R}$ with the Lebesgue measure). Then the space $L^2(X)$ is defined as:

$$L^2(X) = \left\{ f : X \to \mathbb{R} \text{ measurable} \;\middle|\; \int_X |f(x)|^2 \, d\mu(x) < \infty \right\}$$

The space $L^2(X)$ is a Hilbert space with the inner product:

$$\langle f, g \rangle = \int_X f(x) g(x) \, d\mu(x)$$

and the associated norm:

$$\|f\|_{L^2} = \left(\int_X |f(x)|^2 \, d\mu(x) \right)^{1/2}$$

Definition of the ℓ^2 Space (Little ℓ^2)

The space ℓ^2 is the set of all square-summable sequences:

$$\ell^2 = \left\{ (x_n)_{n=1}^{\infty} \;\middle|\; \sum_{n=1}^{\infty} |x_n|^2 < \infty \right\}$$

It is equipped with the inner product:

$$\langle x, y \rangle = \sum_{n=1}^{\infty} x_n y_n$$

and the corresponding norm:

$$\|x\|_{\ell^2} = \left(\sum_{n=1}^{\infty} |x_n|^2 \right)^{1/2}$$

Connection between L^2 and ℓ^2

Let's first define what **orthonormal system** means. Basically, an orthonormal system contains orthogonal vectors (easy to check by its dot product, it should be 0, in other words the 2 vectors are perpendicular) and normalized vectors (their norm is 1). More formally for $(\varphi_n) \subset L^2(X)$

$$\langle \varphi_k, \varphi_j \rangle = \begin{cases} 1 & \text{if } k = j, \\ 0 & \text{if } k \neq j. \end{cases}$$

The spaces L^2 and ℓ^2 are **isometrically isomorphic**, meaning:

1. We can choose a **complete orthonormal system** $(\varphi_n) \subset L^2(X)$ (e.g., Fourier basis).

2. For any $f \in L^2(X)$, let's define its **Fourier coefficients** (since Fourier series can approximate square-integrable functions):

$$c_n = \langle f, \varphi_n \rangle$$

3. Then the map $T : L^2(X) \to \ell^2$, defined by:

$$T(f) = (c_n)_{n=1}^{\infty}$$

is a **linear isometric isomorphism**, since:

$$\|f\|_{L^2}^2 = \sum_{n=1}^{\infty} |c_n|^2 = \|T(f)\|_{\ell^2}^2$$

Linear isometric isomorphism in simpler terms means that the two spaces "behave" identically.

This means L^2 and ℓ^2 have the same structure from the perspective of Hilbert space theory.

10 Functions in Hilbert space

For Hilbert space $L^2([-1, 1])$ if we take $1, x, x^2, ...$ and apply Gram-Schmidt orthogonalization (a process for turning the set of linearly independent vectors into a set of orthonormal vectors), the result will be complete orthonormal system (Figure 6) known as Lebesgue polynomials.

For Hilbert space $L^2([0, \infty))$ if we take $x^n e^{-x}, n = 0, 1, ...$ and apply Gram-Schmidt orthogonalization, the result will be complete orthonormal sequence of Laguerre functions (Figure 1).

In Hilbert space $L^2([0,1])$ one can define wavelets, such as Haar functions (Figure 5).

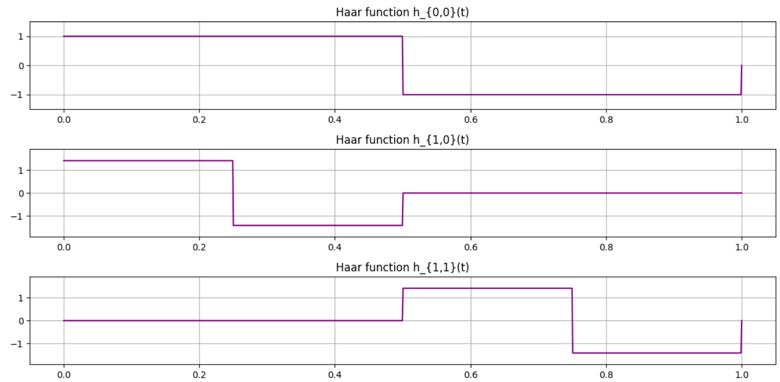


Figure 5: Haar function with different parameters

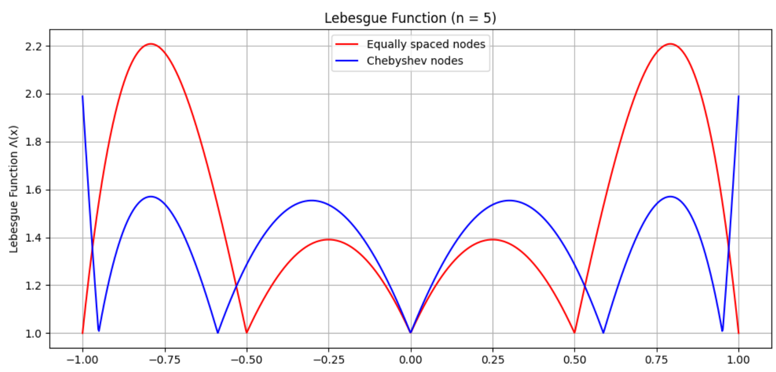


Figure 6: Lebesgue polynomial of degree n

11 Hilbert Spaces in Machine Learning

Hilbert spaces are used in the theory of kernel methods in machine learning. Kernel functions map input data to a space where it might become separable. For instance, kernels are used in support vector machines (SVM) to transform data space from a linearly non-separable into separable. Support vector machines in short, create a hyperplane (plane generalized to higher number of dimensions) to separate data for classifying it.

Reproducing Kernel Hilbert Space (RKHS) is a special type of Hilbert space of functions. Evaluation at a point can be done with an inner product using a kernel.

RKHS enables SVMs to find hyperplanes that separate data points in the transformed feature space, even when the original data is not linearly separable (Figure 7).

RKHS is also used to define covariance functions for probabilistic regression models.

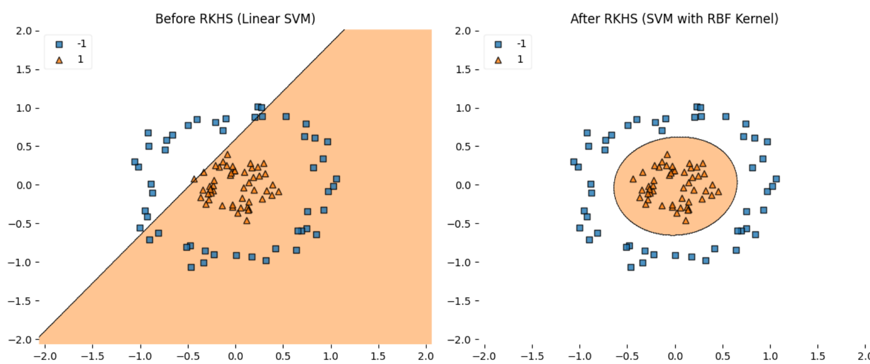


Figure 7: RKHS makes data separable

Quantum Mechanics

Hilbert spaces enable the representation of states, probabilities in quantum mechanics. We won't dive deep in the application of Hilbert space in this field.

12 Conclusion

Hilbert spaces provide a framework for reasoning about environments that otherwise would be hard to explore. Their rigorous yet flexible way of approaching and understanding complex systems enable their application in machine learning, physics, data science, and of course mathematics. This paper focused on the definitions leading up to the Hilbert space and then on examples of spaces in the Hilbert space with their characteristic properties.

Sources

References

[1] Robert G Bartle and Donald R Sherbert. *Introduction to real analysis*. John Wiley & Sons, Inc., 2000.

[2] Christopher M Bishop and Nasser M Nasrabadi. *Pattern recognition and machine learning*. Vol. 4. 4. Springer, 2006.

[3] David Ronald Kincaid and Elliott Ward Cheney. *Numerical analysis: mathematics of scientific computing*. Vol. 2. American Mathematical Soc., 2009.

[4] Erwin Kreyszig. *Introductory functional analysis with applications*. John Wiley & Sons, 1991.

[5] Dan Margalit, Joseph Rabinoff, and Larry Rolen. "Interactive linear algebra". In: *Georgia Institute of Technology* (2017), p. 18.

[6] Karen Saxe. *Beginning functional analysis*. Vol. 16. Springer, 2002.

[7] Gilbert Strang. *Introduction to linear algebra*. SIAM, 2022.