Chapter 2

Function Spaces

Beauty is the first test: there is no permanent place in the world for ugly mathematics.

– G.H. Hardy

FUNCTIONAL spaces are central to analysis and applied mathematics, particularly in formulating and solving differential equations. This chapter introduces some basic properties of functional spaces and discusses several common examples, including Lebesgue spaces and Sobolev spaces. The in-depth understanding can be found in Stein and Shakarchi (2011) and Evans (2022).

2.1 Complete Metric Space

Definition 2.1.1. *A **metric space** is defined by an ordered pair (M, d), where M is a set and d is a metric on a set M, which is defined as a function $d \colon M \times M \mapsto \mathbb{R}^+ \cup \{0\}$ that*

1. $d(x, y) = d(y, x)$ *(symmetry)*
2. $d(x, y) \geq 0$ *and* $d(x, y) = 0$ *iff* $x = y$
3. $d(x, y) + d(y, z) \geq d(x, z)$ *(triangle inequality)*

If the symmetry property is absent, then d is a quasi-metric. If the positive definiteness is lacking, then d is a semi-metric. The metric d describes the distances between elements in M (and automatically defines "open sets" or "topology").

Example 2.1.2. *On Euclidean space $\mathbb{R}^n$, the Euclidean distance*

$$d_2(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_{\ell^2} := \left(\sum_{i=1}^{n} |x_i - y_i|^2 \right)^{1/2}$$

is the most common metric.

It can be shown that $d_p(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_{\ell^p} := \left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{1/p}$ is also a metric when $p \geq 1$. There are other metrics on $\mathbb{R}^d$. For instance, if $d(\mathbf{x}, \mathbf{y})$ is a metric on $\mathbb{R}^n$, then

$$\delta(\mathbf{x}, \mathbf{y}) := \frac{d(\mathbf{x}, \mathbf{y})}{1 + d(\mathbf{x}, \mathbf{y})}$$

is also a metric (see Exercises 2.5.2).

With the concept of **metric**, we can define **convergence** of a sequence concretely.

Definition 2.1.3. *A point $x \in M$ is the limit of a sequence $\{x_n\}_{n \geq 1} \subset M$ if for each $\epsilon > 0$, there exists a natural number N that, for every $n > N$, $d(x_n, x) < \epsilon$.*

Definition 2.1.4. *A sequence $\{x_n\}_{n \geq 1} \subset M$ is a Cauchy sequence if for any $\epsilon > 0$, there exists a natural number N that, for every $m, n > N$, $d(x_n, x_m) < \epsilon$.*

It is fairly simple to show that a convergent sequence is a Cauchy sequence, but the converse is not true.

Example 2.1.5. *For instance, let (M, d) be the set of continuous functions on $[0, 1]$ with an induced L^2 metric*

$$d(f, g) := \int_0^1 |f(t) - g(t)|^2 dt.$$

If the sequence satisfies that

$$x_n(t) = \begin{cases} 1 - nt, & 0 \leq t \leq n^{-1}, \\ 0, & \textit{otherwise}, \end{cases}$$

Then the sequence converges under the given metric

$$\int_0^1 |x_n(t) - x_m(t)|^2 dt \leq 2 \int_0^1 |x_n(t)|^2 + |x_m(t)|^2 dt \to 0.$$

But $\lim_{n\to\infty} x_n$ is not a continuous function.

Regarding applications, many numerical programs find the solutions through sequential improvements until a certain stopping criterion is met. However, it does not imply the sequence can eventually arrive at the exact solution. We introduce the concept of **closeness** in metric spaces.

> **Definition 2.1.6.** *A metric space (M, d) is* ***complete*** *if every* ***Cauchy*** *sequence in M has a limit in M.*

The **completeness** of a metric space is an essential concept in studying the limiting behavior. For instance, the rational numbers $\mathbb{Q} \subset \mathbb{R}$ are not a complete metric space with the Euclidean metric since one can find a sequence of rational numbers approaching an irrational number. A non-complete metric space could be "completed" if another metric is used. In the previous example 2.1.5, if the metric is replaced by

$$d(f, g) := \max_{x\in[0,1]} |f - g|,$$

then every Cauchy sequence of continuous functions on $[0, 1]$ must converge to a continuous function (see Exercises 2.5.3).

2.2 Banach Spaces

Definition 2.2.1. *A vector space V is called **normed space** if there is a norm function $\|\cdot\|: V \mapsto \mathbb{R}^+ \cup \{0\}$ that*

1. $\|\mathbf{x}\| \geq 0$ *and* $\|\mathbf{x}\| = 0$ *iff* $\mathbf{x} = \mathbf{0}$.
2. $\|a\mathbf{x}\| = |a| \cdot \|\mathbf{x}\|$.
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

There a norm is given on V, then $d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|$ is an induced metric.

Example 2.2.2. *The ℓ^p space consists of the sequence $\{a_k\}_{k\geq 1}$ that*

$$\sum_{k\geq 1} |a_k|^p < \infty.$$

When $p \geq 1$, the ℓ^p space is a Banach space. The norm of $x := \{x_k\}_{k\geq 1}$ is naturally defined by $\|x\| = \left(\sum_{k\geq 1} |x_k|^p\right)^{1/p}$. The triangle inequality of $\|\cdot\|$ uses the convexity of the function $f : t \mapsto |t|^p$ (prove it). Let $s = \frac{\|x\|}{\|x\|+\|y\|}$, then by the convexity,

$$\left| s\frac{x_k}{\|x\|} + (1-s)\frac{y_k}{\|y\|} \right|^p \geq s\frac{|x_k|^p}{\|x\|^p} + (1-s)\frac{|y_k|^p}{\|y\|^p}$$

Taking the sum over k, it implies

$$\sum_{k\geq 1} \left| \frac{x_k + y_k}{\|x\| + \|y\|} \right|^p \geq 1.$$

Example 2.2.3. *Let $C([0,1])$ be the vector space of continuous functions on $[0,1]$. Then, it is a normed space with one of the following norm functions*

1. $\|f\|_\infty := \sup_{[0,1]} |f(x)|$ *(prove it).*
2. $\|f\|_{L^2} := \left(\int_0^1 |f(x)|^2 dx \right)^{1/2}$. *The triangle inequality can be proved similarly to Example 2.2.2.*

However, we already know $C[0,1]$ is not complete under L^2 norm.

Definition 2.2.4. *A complete normed space is called a **Banach space**.*

Example 2.2.5. *Let $C^1([0,1])$ be the vector space of functions with continuous first* *derivatives on* $[0,1]$*. It is* not *complete under* $\|\cdot\|_\infty$ *norm. For instance, we can construct a Cauchy sequence* $f_n(x) = \sqrt{(x-\frac{1}{2})^2 + \frac{1}{n}}$*,* $n = 0, 1, \cdots$*,*

$$|f_n - f_m| = \left|\sqrt{(x-\frac{1}{2})^2 + \frac{1}{n}} - \sqrt{(x-\frac{1}{2})^2 + \frac{1}{m}}\right|$$

$$= \left|\frac{\frac{1}{n} - \frac{1}{m}}{\sqrt{(x-\frac{1}{2})^2 + \frac{1}{m}} + \sqrt{(x-\frac{1}{2})^2 + \frac{1}{n}}}\right| < \left|\frac{1}{\sqrt{m}} - \frac{1}{\sqrt{n}}\right| < \varepsilon,$$

if $m, n > \varepsilon^{-2}$*. However, the limit* $f(x) = |x - \frac{1}{2}| \notin \mathcal{C}^1([0,1])$*.*

Example 2.2.6. *The space* $\mathcal{C}([0,1])$ *is not complete under norm*

$$\|f\|_{L^p} := \left(\int_0^1 |f(x)|^p dx\right)^{1/p}$$

for any $p \geq 1$*, since we can find a Cauchy sequence (in* $\|\cdot\|_p$*) of continuous functions converging to a discontinuous function.*

2.3 Lebesgue Spaces

The Lebesgue spaces, which are also known as L^p spaces, are particularly useful in analysis and differential equations. Since the measure theory is *not* a prerequisite for this note, we only give a short tour of the core idea of this class of spaces.

2.3.1 A Short Tour to Lebesgue Integral

As demonstrated in previous examples, the vector space of continuous functions on a compact domain cannot be *complete* with the L^p norm since discontinuities may exist for Cauchy sequences. The purpose of Lebesgue space is to *complete* the space by adding the missing elements (pointwise limits of sequences, if they exist).

However, this will allow functions with mild singularities (e.g., square integrable), such as $f(x) = |x|^{-1/3}$ near the origin. Therefore, the Riemann integral cannot be used to define the L^p norm. We need to extend the definition of Riemann integral to *Lebesgue integral*. In the following context, we focus solely on real-valued functions; there is no technical difficulty in generalizing to complex-valued functions.

Definition 2.3.1. *Let $K(\mathbb{R})$ be the set of real-valued continuous functions with compact support in $\mathbb{R}$.*

The Riemann integral is well-defined for $K(\mathbb{R})$. We denote the Riemann integral of $f \in K(\mathbb{R})$ by $I(f)$. To eliminate the difficulty of the Riemann integral for unbounded functions, we introduce the following "integral" $V(f)$.

Definition 2.3.2. *Let f be a real-valued function on $\mathbb{R}$. Define*

$$V(f) := \inf\left\{ \sum_{k=1}^{\infty} I(f_k) \colon f_n \in K(\mathbb{R}), f_n \geq 0, |f(x)| \leq \sum_{k=1}^{\infty} f_n(x)\,, \forall x \in \mathbb{R} \right\}$$

The definition of $V(f)$ is well-defined even if f can attain infinity. One thing we can immediately find out is $V(f) = I(|f|)$ if $f \in K(\mathbb{R})$ and $V(\alpha f) = \alpha V(f)$. We can also deduce the triangle inequality (prove it).

Lemma 2.3.3. *Let f and f_k be real-valued functions on $\mathbb{R}$. If $|f(x)| \leq \sum_{k=1}^{\infty} f_n(t)$, then*

$$V(f) \leq \sum_{k=1}^{\infty} V(f_n).$$

The function V satisfies the triangle inequality and homogeneity, but V does not have to be positive definite. Thus, V is only a semi-norm (but $I(|f|)$ is a norm when $f \in K(\mathbb{R})$).

Definition 2.3.4. *Let the Lebesgue space $L^1(\mathbb{R})$ be the set of functions $f : \mathbb{R} \to \mathbb{R}$ such that for any $\varepsilon > 0$ there exists a $g \in K(\mathbb{R})$ that $V(f-g) < \varepsilon$. Define*

$$I'(f) := \lim_{n\to\infty} I(f_n), \quad f \in L^1(\mathbb{R}),$$

where $f_n \in K(\mathbb{R})$ and $V(f - f_n) \to 0$.

Note that this definition does not involve measure theory. First, we can see $L^1(\mathbb{R})$ is a linear vector space that contains $K(\mathbb{R})$. Second, we need to show $I'(f)$ is well-defined for $f \in L^1(\mathbb{R})$. Since for sufficiently large n and m,

$$|I(f_n) - I(f_m)| \le I(|f_n - f_m|) = V(f_n - f_m) \le V(f_n - f) + V(f_m - f) \to 0.$$

The triangle inequality and the homogeneity of I' are inherited from I. We can see that the definition of $L^1(\mathbb{R})$ does *not* distinguish two functions if their difference vanishes under the semi-norm V.

Theorem 2.3.5 (Completeness). *If $\{f_n\}_{n\ge1}$ is a Cauchy sequence in $L^1(\mathbb{R})$ with the semi-norm V, then there exists $f \in L^1(\mathbb{R})$ that $V(f - f_n) \to 0$ as $n \to \infty$. That is, $L^1(\mathbb{R})$ equipped with semi-norm $V(f)$ is complete.*

Proof. Since $\{f_n\}_{n\ge1}$ is a Cauchy sequence, we may assume (up to a subsequence)

$$V(f_n - f_{n+1}) < 2^{-n}.$$

Let $f(x) = \limsup_{n\to\infty} f_n(x)$ for each $x \in \mathbb{R}$, it is possible f attains infinity at certain points. Then (why?)

$$|f(x) - f_n(x)| \le \sum_{k=n}^{\infty} |f_{k+1}(x) - f_k(x)|, \quad \forall x \in \mathbb{R}.$$

This inequality holds no matter whether the $f(x)$ is infinite or finite. The semi-norm's triangle inequality leads to

$$V(f - f_n) \le \sum_{k=n}^{\infty} V(f_{k+1} - f_k) \le 2^{-(n-1)}.$$

Therefore $f \in L^1(\mathbb{R})$. Thus, $L^1(\mathbb{R})$ is a Banach space. □

Theorem 2.3.6 (Monotone Convergence Theorem). *Let $f_n \in L^1(\mathbb{R})$ and $0 \le f_1 \le f_2 \le \cdots$ such that $\lim_{n\to\infty} I'(f_n) < \infty$. If $f(x) = \lim_{n\to\infty} f_n(x)$ exists for all $x \in \mathbb{R}$, then $f \in L^1(\mathbb{R})$ and*

$$I'(f) = \lim_{n\to\infty} I'(f_n).$$

Proof. Apply the triangle inequality for the semi-norm V to the following inequality,

$$|f(x) - f_n(x)| \leq \sum_{k=n}^{\infty} (f_{k+1}(x) - f_k(x)),$$

then

$$\begin{aligned} V(f - f_n) &\leq \sum_{k=n}^{\infty} V(f_{k+1} - f_k) = \sum_{k=n}^{\infty} I'(f_{k+1} - f_k) \\ &= \sum_{k=n}^{\infty} I'(f_{k+1}) - I'(f_k) = \left[\lim_{k\to\infty} I'(f_k)\right] - I'(f_n). \end{aligned}$$

Therefore, $I'(f) = \lim_{n\to\infty} I'(f_n)$. □

Corollary 2.3.7 (Dominated Convergence Theorem). *Suppose $f_n \in L^1(\mathbb{R})$ and there exists $g \in L^1(\mathbb{R})$ that $|f_n(x)| \leq g(x)$ for all $x \in \mathbb{R}$. If $f(x) = \lim_{n\to\infty} f_n(x)$ exists for all $x \in \mathbb{R}$, then $f \in L^1(\mathbb{R})$ and*

$$I'(f) = \lim_{n\to\infty} I'(f_n).$$

Proof. Apply the Monotone Convergence Theorem to the (increasing) sequences $\{g + \inf_{n\leq k} f_k\}_{n\geq 1}$ and $\{g - \sup_{n\leq k} f_k\}_{n\geq 1}$, we find that

$$\limsup_{n\to\infty} f_n = \liminf_{n\to\infty} f_n = f \in L^1(\mathbb{R}).$$

Meanwhile, we have

$$I'(f) = I'(\limsup_{n\to\infty} f_n) = \lim_{n\to\infty} I'(\sup_{n\leq k} f_k) \geq \lim_{n\to\infty} \sup_{n\leq k} I'(f_k) = \limsup_{n\to\infty} I'(f_n),$$

and

$$I'(f) = I'(\liminf_{n\to\infty} f_n) = \lim_{n\to\infty} I'(\inf_{n\leq k} f_k) \leq \lim_{n\to\infty} \inf_{n\leq k} I'(f_k) = \liminf_{n\to\infty} I'(f_n).$$

□

Example 2.3.8. *In general, it is incorrect that $f = \lim_{n\to\infty} f_n$ leads to $I'(f_n) \to I'(f)$ without further assumptions, as in the previous corollaries. For example, let $f_n = \mathbf{1}_{[n,n+1]}$ converges pointwise to constant zero while $I'(f_n) = 1$.*

The $L^1(\mathbb{R})$ space can be easily extended to $L^p(\mathbb{R})$, $p \geq 1$ and higher dimensions. Note that $L^p(\mathbb{R})$ does not contain another $L^{p'}(\mathbb{R})$ as its subspace when $p < p'$. This is different when the Lebesgue spaces are restricted to a domain X

of finite volume, that is, $\mathbf{1}_X \in L^1(\mathbb{R})$. The semi-norm V can be viewed as a norm modulo the equivalence class $\{f \in L^1(\mathbb{R})\colon V(f) = 0\}$. Then, we can define the usual L^1 norm by $\|f\|_1 = V(f)$, and $\|f\|_p = \|f^p\|_1^{1/p}$.

Example 2.3.9. *The function* $(1+|x|)^{-1} \in L^2(\mathbb{R})$ *but not in* $L^1(\mathbb{R})$*. The function* $|x|^{-1/2}/(1+|x|^2) \in L^1(\mathbb{R})$ *but not in* $L^2(\mathbb{R})$*.*

The following Hölder inequality is essential for Lebesgue space.

Theorem 2.3.10 (Hölder's inequality)**.** *Suppose that* $1 \le p \le \infty$ *and* $1 < q < \infty$ *that* $p^{-1}+q^{-1} = 1$*. If* $f \in L^p(\mathbb{R})$ *and* $g \in L^q(\mathbb{R})$*, then* $fg \in L^1(\mathbb{R})$*. Moreover,*

$$\|f\|_{L^p}\|g\|_{L^q} \ge \|fg\|_{L^1}.$$

Proof. The basic tool is the following inequality

$$x^a \le (1-a) + ax, \quad a \in (0,1),\ x \ge 0.$$

This is due to the convexity of the function $h(a) = x^a$, since $h''(a) = (\log x)^2 x^a \ge 0$. Then it implies that

$$x^a y^{1-a} \le ax + (1-a)y, \quad a \in (0,1),\ x, y \ge 0.$$

Set $x = \frac{|f|^p}{\|f\|_{L^p}^p}$, $y = \frac{|g|}{\|g\|_{L^q}^q}$ and $a = \frac{1}{p}$, then

$$\left[\frac{|f|^p}{\|f\|_{L^p}^p}\right]^{1/p}\left[\frac{|g|^q}{\|g\|_{L^q}^q}\right]^{1/q} \le \frac{1}{p}\frac{|f|^p}{\|f\|_{L^p}^p} + \frac{1}{q}\frac{|g|}{\|g\|_{L^q}^q}.$$

Integration on both sides yields the inequality. □

The above Hölder inequality implies that the L^p norm is truly a norm (the idea in Example 2.2.2 also applies).

Corollary 2.3.11 (Minkowski's inequality)**.** *Let* $1 \le p \le \infty$*, then*

$$\|f+g\|_{L^p} \le \|f\|_{L^p} + \|g\|_{L^p}.$$

Proof. Note that

$$\int_{\mathbb{R}} |f+g|^p dx \le \int_{\mathbb{R}} |f||f+g|^{p-1} dx + \int_{\mathbb{R}} |g||f+g|^{p-1} dx.$$

Then apply Hölder's inequality for both terms,

$$\int_{\mathbb{R}} |f||f+g|^{p-1}dx \le \|f\|_{L^p}\||f+g|^{p-1}\|_{L^q} = \|f\|_{L^p}\cdot\|f+g\|_{L^p}^{p/q}$$
$$\int_{\mathbb{R}} |g||f+g|^{p-1}dx \le \|g\|_{L^p}\||f+g|^{p-1}\|_{L^q} = \|g\|_{L^p}\cdot\|f+g\|_{L^p}^{p/q}$$

Thus,

$$\int_{\mathbb{R}} |f+g|^p dx \le (\|f\|_{L^p} + \|g\|_{L^p})\cdot\|f+g\|_{L^p}^{p/q}.$$

□

2.3.2 Interpolation of Lebesgue Spaces

The interpolation is a fundamental tool in applied analysis. Unlike the qualitative properties (boundedness, convergence, etc.), the interpolation technique often provides quantitative estimates. For instance, if f satisfies certain properties $\mathcal{A}_0$ and $\mathcal{A}_1$, then we may conclude a family of similar properties $\mathcal{A}_\theta$ for all $\theta \in (0,1)$. The following theorem is a particular example.

Theorem 2.3.12 (Interpolation). *Let $p_0 < p_1$ and suppose $f \in L^{p_0}(\mathbb{R})$ and $f \in L^{p_1}(\mathbb{R})$, then $f \in L^p(\mathbb{R})$ for any $p \in (p_0, p_1)$. Moreover,*

$$\|f\|_{p_0}^{\alpha}\|f\|_{p_1}^{1-\alpha} \ge \|f\|_p.$$

Proof. It suffices to prove for $f \in K(\mathbb{R})$ (why?). Then let $\frac{1}{p} = \frac{\alpha}{p_0} + \frac{(1-\alpha)}{p_1}$, $\alpha \in (0,1)$ and apply Hölder inequality,

$$\left(\int_{\mathbb{R}} |f^{p_0}(x)|dx\right)^{\alpha p/p_0}\left(\int_{\mathbb{R}} |f^{p_1}(x)|dx\right)^{(1-\alpha)p/p_1} \ge \int_{\mathbb{R}} |f^p(x)|dx.$$

□

Example 2.3.13. *If a function $f \in L^p(\mathbb{R}) \cap L^q(\mathbb{R})$ that $1 \le p < q < \infty$, then $f \in L^p(\mathbb{R}) + L^q(\mathbb{R})$ due to the decomposition*

$$f(x) = f(x)\mathbf{1}_{|f|>1} + f(x)\mathbf{1}_{|f|\le 1}.$$

The interpolation theorem can be interpreted as

$$\alpha \log \|f\|_{p_0} + (1-\alpha)\log\|f\|_{p_1} \ge \log\|f\|_p,$$

This is known as the *log-convexity* for the mapping $\frac{1}{p} \to \log \|f\|_p$.

2.4 Approximation in Hilbert Spaces

Similar to linear vector spaces (of finite dimension), we can define an inner product on function spaces (of infinite dimension).

Definition 2.4.1. *Let S be a function space. The inner product $\langle \cdot, \cdot \rangle : S \times S \to \mathbb{R}$ (or $\to \mathbb{C}$) satisfies*

1. $\langle f, g \rangle = \overline{\langle g, f \rangle}$.
2. $\langle \alpha f, g \rangle = \alpha \langle f, g \rangle$
3. $\langle f + g, h \rangle = \langle f, h \rangle + \langle g, h \rangle$
4. $\langle f, f \rangle \geq 0$*, the equal sign occurs only if* $f \equiv 0$.

The inner product necessarily defines a norm on the function space, thereby creating a normed space.

Definition 2.4.2. *A complete inner product space is called a* ***Hilbert*** *space.*

A Hilbert space is also a Banach space, but with an additional inner-product structure. The most popular Hilbert space is the specific Sobolev space $H^k(\mathbb{R})$ (can be extended to higher dimensions easily), which is defined by

$$H^k(\mathbb{R}) := \{f \in L^2(\Omega) \mid D^i f \in L^2(\mathbb{R}),\ i = 0, 1, \cdots, k\},$$

where D denotes the **weak derivative** instead of a true derivative. The inner product is defined as

$$\langle f, g \rangle_{H^k} = \sum_{i=0}^{k} \langle D^i f, D^i g \rangle_{L^2}.$$

The weak derivatives become the true derivatives when $f \in C_c^k(\mathbb{R})$, but the norm

$$\|f\| = \left(\sum_{i=0}^{k} \int_{\mathbb{R}} |\partial^i f(x)|^2 dx \right)^{1/2}$$

does not make the space $C_c^k(\mathbb{R})$ complete. One can view H^k as the completion of

$C_c^k(\mathbb{R})$ in the above norm. It automatically induces the inner product structure. The weak derivatives will be covered systematically in a later chapter. In the following, we focus on the abstract Hilbert space, and the norm is inherited from its inner product.

2.4.1 Generalized Fourier Series

Let H be a general metric space, and $f \notin C$, where C is a subset of H. It is common to seek a function in C that can approximate f. This is usually difficult and related to explicit but heuristic construction. However, if H is a Hilbert space and the approximation is measured under its norm, there is a systematic (and geometric) way to solve. The problem can be formulated as

$$h^* = \arg\min_{h \in C} \|f - h\|^2. \tag{2.1}$$

There are two essential questions to study: whether h^* exists and whether h^* is unique.

Theorem 2.4.3. *If H is a Hilbert space and $C \subset H$ is closed and convex, then $h^* \in C$ in (2.1) exists and is unique.*

Proof. Let h_k be a minimizing sequence such that

$$\|f - h_k\|^2 < \inf_{h \in H} \|f - h\|^2 + \frac{1}{k}$$

By convexity, $\frac{h_n + h_m}{2} \in C$, then

$$\begin{aligned} \|h_m - h_n\|^2 &= 2\|h_m - f\|^2 + 2\|h_n - f\|^2 - 4\left\|\frac{h_n + h_m}{2} - f\right\|^2 \\ &\le 2(\inf_{h \in C} \|f - h\|^2 + \frac{1}{m}) + 2(\inf_{C \in H} \|f - h\|^2 + \frac{1}{n}) - 4\inf_{h \in C} \|f - h\|^2 \\ &= \frac{2}{m} + \frac{2}{n}. \end{aligned}$$

This means that h_k forms a Cauchy sequence and converges to a limit in C because C is closed in H.

If h^* and h^{**} both minimizes (2.1), then

$$\|h^* - h^{**}\|^2 = 2\|h^* - f\|^2 + 2\|h^{**} - f\|^2 - 4\left\|\frac{h^* + h^{**}}{2} - f\right\|^2$$

If $h^* \neq h^{**}$, it means $\frac{h^*+h^{**}}{2}$ achieves a smaller error, contradiction. $\square$

Moreover, we can see that $f - h^* \perp C$, that is, any element $x \in C$ must satisfy $\langle x, f - h^* \rangle = 0$. This can be geometrically understood as h^* is the *orthogonal projection* onto C.

When C is a subspace spanned by a basis $\{\phi_i\}_{i=1}^N \subset H$, then there exists a unique combination

$$h^* = \sum_{i=1}^{N} \alpha_i \phi_i$$

for the minimizer of (2.1). In addition, using the orthogonal relation $h^* - f \perp C$, we have $\langle f - h^*, \phi_i \rangle = 0$. In particular, if $\{\phi_i\}_{i=1}^N$ is an orthonormal basis, then

$$\langle f, \phi_i \rangle = \alpha_i \langle \phi_i, \phi_i \rangle = \alpha_i.$$

Therefore, the minimizer

$$h^* = \sum_{i=1}^{N} \langle f, \phi_i \rangle \phi_i.$$

This is also called a *generalized Fourier series*, and

$$\|f - h^*\|^2 = \|f\|^2 - \sum_{i=1}^{N} |\langle f, \phi_i \rangle|^2.$$

In other words, $\sum_{i=1}^N |\langle f, \phi_i \rangle|^2 \leq \|f\|^2$. This is known as Bessel's inequality; it holds whether N is finite or not. However, an infinite orthonormal basis does not imply $h^* = f$ eventually. For instance,

$$C = \text{span}\{\sin nx\}_{n=1}^{\infty}$$

forms a subspace of Hilbert space $H = L^2[0, 2\pi]$. However, $C \neq H$ since H contains functions like $\cos nx$ which is orthogonal to C.

Definition 2.4.4. *An orthonormal basis $\{\phi_i\}_{i=1}^{\infty}$ is complete if and only if*

$$f = \sum_{i=1}^{\infty} \langle f, \phi_i \rangle \phi_i$$

for all $f \in H$.

Therefore, we must have

$$\|f\|^2 = \sum_{i=1}^{\infty} |\langle f, \phi_i\rangle|^2.$$

This is called *Parseval's identity*. We show a few examples of a complete basis.

Example 2.4.5. *The classical complete Fourier basis (unnormalized) for $L^2[0, 2\pi]$ is*

$$\{1, \sin x, \cos x, \sin 2x, \cos 2x, \cdots\}.$$

Example 2.4.6. *Let $H = \{-1, 1\}^n$ be the set of hypercube vertices. It has 2^n vertices in total, each vertex has coordinate $(x_1, x_2, \cdots, x_n)$, and $x_i \in \{-1, 1\}$. We can define $L^2(H)$ through functions $f : H \to \mathbb{R}$ such that*

$$\sum_{x \in H} |f(x)|^2 < \infty.$$

It is straightforward to see that $L^2(H)$ is a Hilbert space and has dimension 2^n. The (unnormalized) basis functions are

$$\phi_S = \prod_{i \in S \subset \{1,2,\cdots,n\}} x_i, \quad S \subset \{1, 2, \cdots, n\}.$$

The special case $\phi_\emptyset = 1$. It is not difficult to check the orthogonality

$$\sum_{x \in H} \phi_S(x)\phi_T(x) = 0, \quad S \neq T.$$

2.4.2 Orthogonal Polynomials

The orthogonal polynomials are particular sets of bases for common function spaces. The classical Weierstrass Theorem implies that the polynomial space is dense in the space of continuous functions (compactly supported), hence dense in any Lebesgue space $L^p([a, b])$. In particular, if we introduce a weight function $\omega(x)$, the inner product

$$\langle f, g\rangle := \int_a^b \omega(x)f(x)g(x)dx \tag{2.2}$$

induces a Hilbert space H such that $\sqrt{\omega} f \in L^2([a, b])$ if and only if $f \in H$. To find a complete orthonormal polynomial basis for H, we introduce the recursion formula.

Definition 2.4.7. *Let Π_n be the set of polynomials of degree n with non-degeneracy, that is,*

$$f^{(n)} \neq 0, \quad \forall f \in \Pi_n.$$

The recursion formula aims to find the orthogonal polynomials $p_n \in \Pi_n$, $n \geq 0$, with the inner product (2.2). First, we can represent (without normalization constant)

$$p_n(x) = xp_{n-1}(x) + \sum_{j=0}^{n-1} \alpha_j p_j(x).$$

The orthogonality relation implies

$$0 = \langle p_n, p_{n-1} \rangle = \langle xp_{n-1}, p_{n-1} \rangle + \alpha_{n-1} \langle p_{n-1}, p_{n-1} \rangle.$$

For $j \leq n-3$, we have

$$0 = \langle p_n, p_j \rangle = \langle p_{n-1}, xp_j \rangle + \alpha_j \langle p_j, p_j \rangle = \alpha_j \langle p_j, p_j \rangle.$$

Therefore, $\alpha_j = 0$ for all $j \leq n-3$. The recursion formula can be reduced to

$$p_n(x) = (x - \beta_n) p_{n-1} + \alpha_{n-1} p_{n-2}.$$

The two constants can be determined by

$$\beta_n = \frac{\langle xp_{n-1}, p_{n-1} \rangle}{\langle p_{n-1}, p_{n-1} \rangle}, \quad \alpha_{n-1} = \frac{\langle p_{n-1}, xp_{n-2} \rangle}{\langle p_{n-2}, p_{n-2} \rangle} = \frac{\langle p_{n-1}, p_{n-1} \rangle}{\langle p_{n-2}, p_{n-2} \rangle}$$

There are a few common sets of orthogonal polynomials.

Example 2.4.8. *Weight function $\omega(x) = \frac{1}{\sqrt{1-x^2}}$ on $[-1, 1]$. The polynomials are*

$$p_n(x) = \cos(n \arccos(x)),$$

called Chebyshev polynomials of the first kind. The recursive formula is

$$T_n(x) = 2xT_{n-1} - T_{n-2}(x), \quad T_0 = 1, \ T_1 = x.$$

Example 2.4.9. *Weight function $\omega(x) = \sqrt{1-x^2}$ on $[-1, 1]$. The polynomials are*

$$p_n(x) = \frac{\sin((n+1)\arccos(x))}{\sin(\arccos x)},$$

called Chebyshev polynomials of the second kind. The recursive formula is

$$U_n(x) = 2xU_{n-1} - U_{n-2}(x), \quad U_0 = 1, \ U_1 = 2x.$$

Remark 2.4.10. *The above two kinds of Chebyshev polynomials are closely related to trigonometric modes* $\cos(n\theta)$ *and* $\sin(n\theta)$ *after a change of variable* $\theta = \arccos x$, $\theta \in [0, 2\pi]$. *The following are the famous related theories.*

Theorem 2.4.11 (Fourier Convergence Theorem). *If f is piecewise C^1 continuous on $[0, 2\pi]$, then its Fourier series converges to $\frac{1}{2}[f(x^+) + f(x^-)]$ for each $x \in [0, 2\pi]$.*

Theorem 2.4.12 (Carleson's Theorem, Carleson (1966)). *If $f \in L^p[0, 2\pi]$ for some $p \in (1, \infty]$, then its Fourier series converges to f almost everywhere.*

The proofs of the above theorems are beyond the scope of this note.

Example 2.4.13. *Weight function $\omega(x) = 1$ on $[-1, 1]$. The polynomials are $p_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n}(x^2 - 1)^n$, called Legendre polynomials. The recursive formula is*

$$P_n(x) = \frac{2n-1}{n} x P_{n-1} - \frac{n-1}{n} P_n, \quad P_0 = 1, \ P_1 = x.$$

More general cases like Jacobi polynomials with weight $\omega(x) = (1-x)^\beta (1+x)^\gamma$ can be found in Szegö (1939).

2.5 Exercises

☞ **Problem 2.5.1.** *Let d_i be a (quasi-)metric on the non-empty sets $\{X_i\}_{i=1}^m$. Prove that*

$$d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{m} \frac{1}{2^i} d_i(x_i, y_i), \quad \mathbf{x} = (x_1, \cdots, x_m), \quad \mathbf{y} = (y_1, \cdots, y_m)$$

is a (quasi-)metric on the product set $\mathcal{X} = X_1 \times X_2 \cdots \times X_m$.

☞ **Problem 2.5.2.** *Suppose $d(\mathbf{x}, \mathbf{y})$ is a (quasi-)metric on $\mathbb{R}^n$. Prove that*

$$\delta(\mathbf{x}, \mathbf{y}) := \frac{d(\mathbf{x}, \mathbf{y})}{1 + d(\mathbf{x}, \mathbf{y})}$$

is also a (quasi-)metric on $\mathbb{R}^n$*.*

☞ **Problem 2.5.3.** *Let* M *be the set of continuous functions on* $[0,1]$ *and*

$$d(f,g) := \max_{x\in[0,1]} |f-g|.$$

Prove (M,d) *is a complete metric space.*

☞ **Problem 2.5.4.** *Assume that* $\log|f| \in L^1(\mathbb{R})$ *and* $f \in L^1(\mathbb{R})$*. Show that*

$$\lim_{p\to 0^+} \|f\|_p = \exp\left(\int_{\mathbb{R}} \log|f| dx\right)$$

☕ **Problem 2.5.5.** *Suppose* $\{\phi_i\}_{i=1}^n$ *is a complete orthonormal set of basis for a Hilbert space* H*. Let* $\{\psi_i\}_{i=1}^n$ *be strongly linearly independent elements in* H*, that is,*

$$\psi_k \notin \overline{\text{span}\{\psi_i\}_{i\neq k}}.$$

Show that if $\sum_{i=1}^\infty \|\phi_i - \psi_i\|^2 < \infty$*, then* $\{\psi_i\}_{i=1}^n$ *spans* H*.*

☕ **Problem 2.5.6.** *If* $p_n \in \Pi_n$*,* $n \geq 0$ *forms an orthogonal polynominal set on* $[a,b]$ *for a weight* $\omega(x) > 0$ *on* (a,b)*, then* p_n *has exactly* n *real roots.*

☕ **Problem 2.5.7.** *Let* $P_n(x)$ *be the Legendre polynomial of degree* n*. Show that*

$$\int_{-1}^{1} |P_n(x)|^2 dx = \frac{2}{2n+1}.$$

Extended Reading

Carleson, L. (1966). On convergence and growth of partial sums of fourier series.

Evans, L. C. (2022). *Partial differential equations*, volume 19. American mathematical society.

Stein, E. M. and Shakarchi, R. (2011). *Functional analysis: introduction to further topics in analysis*, volume 4. Princeton University Press.

Szegö, G. (1939). *Orthogonal polynomials*, volume 23. American Mathematical Soc.