Exploration of Orthogonal Polynomials

Nidhruv Murali

June 2025

Abstract

Orthogonal polynomials arise naturally from the abstraction of vector spaces and inner products into the realm of functions and algebraic expressions. This paper explores the foundational ideas that allow polynomials to behave like vectors, enabling the powerful machinery of linear algebra to be applied to problems in approximation theory, quantum mechanics, and numerical analysis. Beginning with an overview of vector and inner product spaces, we define orthogonal polynomials and construct them using the Gram-Schmidt process. We then investigate classical families of orthogonal polynomials, including Legendre, Chebyshev, Hermite, and Laguerre, and examine their recurrence relations, differential properties, and practical applications. Special attention is given to how these polynomials approximate functions more effectively than Taylor series across intervals, and how their root structures lend themselves to elegant and efficient numerical algorithms. Ultimately, this paper highlights how orthogonal polynomials form a unifying thread between pure mathematical theory and practical computational tools.

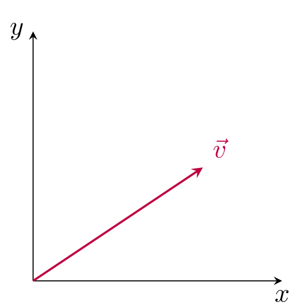
1 Introduction

Abstraction is a foundational theme in mathematics. In linear algebra, we learn to see vectors not just as arrows in space, but as abstract objects governed by algebraic structure. This perspective opens the door to studying functions, polynomials, and other mathematical entities using linear techniques.

Orthogonal polynomials emerge naturally when we apply inner product structure to the space of polynomials. They generalize the geometric idea of perpendicular vectors to the function world, and they play a central role in areas like approximation theory, numerical methods, and quantum mechanics. This paper develops the theory of orthogonal polynomials from the ground up, starting with vector spaces and inner products, and builds toward their applications and structure.

2 Vector Spaces

In a typical linear algebra class, one of the first concepts you encounter is the vector. Initially, vectors are introduced geometrically as arrows that have both direction and magnitude. In two dimensions, this visual approach helps build intuition about operations like vector addition and scalar multiplication.



After seeing vectors as arrows in space, one of the first properties we study is their **length**. In two dimensions, the length of a vector $\vec{v} = \langle v_1, v_2 \rangle$ is given by the Pythagorean theorem:

$$\|\vec{v}\| = \sqrt{v_1^2 + v_2^2}$$

As we move beyond two or three dimensions, we need more general tools. To do this, we define an operation between vectors called the **dot product**, which gives us both geometric and algebraic insight:

$$\vec{v} \cdot \vec{w} = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n$$

Using this definition, we can compute the length (or *norm*) of a vector in any dimension:

$$\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}}$$

The dot product also allows us to define when two vectors are **orthogonal**: two vectors are perpendicular if and only if their dot product is zero.

This idea of an inner multiplication that reflects geometric structure becomes crucial as we transition from finite-dimensional vectors to functions, polynomials, and beyond.

As we continue to explore vectors, a natural question arises: what actually *is* a vector? In two or three dimensions, we think of arrows. But in mathematics, we often follow a general philosophy of **abstraction**: we extract the essential properties of an object and generalize them into a broader definition.

In the case of vectors, this leads us to the idea of a **vector space**. Instead of thinking only about arrows in space, we define a set along with rules for addition and scalar multiplication that mimic the behavior of familiar vectors.

Definition: Vector Space A **vector space** V over a field $\mathbb{F}$ (like $\mathbb{R}$ or $\mathbb{C}$) is a set equipped with:

- A vector addition operation $+ : V \times V \to V$
- A scalar multiplication operation $\cdot : \mathbb{F} \times V \to V$

such that the following axioms hold for all $\vec{u}, \vec{v}, \vec{w} \in V$ and all scalars $a, b \in \mathbb{F}$:

1. $\vec{u} + \vec{v} = \vec{v} + \vec{u}$ (commutativity)
2. $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$ (associativity)
3. There exists a zero vector $\vec{0} \in V$ such that $\vec{v} + \vec{0} = \vec{v}$
4. Every $\vec{v} \in V$ has an additive inverse $-\vec{v} \in V$ such that $\vec{v} + (-\vec{v}) = \vec{0}$
5. $a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v}$ (scalar distributivity)
6. $(a + b)\vec{v} = a\vec{v} + b\vec{v}$
7. $a(b\vec{v}) = (ab)\vec{v}$
8. $1 \cdot \vec{v} = \vec{v}$

These axioms may seem vague, but that's exactly the point. They are abstract enough to apply not just to arrows in space, but also to objects like polynomials, functions, matrices, and more. As long as a set supports addition and scalar multiplication in a way that satisfies these rules, it can be considered a vector space.

One of the most surprising and beautiful aspects of abstraction in linear algebra is that many seemingly unrelated mathematical objects turn out to behave like vectors. Because they satisfy the same set of vector space axioms, we can treat them analogously. Below are several important examples of vector spaces:

- **Euclidean space** $\mathbb{R}^n$: This is the most familiar example, consisting of ordered n-tuples of real numbers. Vector addition and scalar multiplication are defined component-wise. This space underlies most of introductory linear algebra.
- **Complex space** $\mathbb{C}^n$: Similar to $\mathbb{R}^n$, but with complex-valued components. Scalar multiplication is done using complex numbers. This vector space is important in electrical engineering, quantum mechanics, and signal processing.
- **Function spaces**: The set of all real-valued functions defined on an interval $[a, b]$ forms a vector space. The sum of two functions is a function, and scalar multiples are defined pointwise. This example is surprising at first — we can treat entire functions as vectors.
- **Polynomial space** $\mathbb{P}_n$: The set of all polynomials of degree at most n, with real coefficients, forms a vector space. The operations are defined by adding polynomials and multiplying them by scalars term-by-term.
- **Matrices** $M_{m\times n}(\mathbb{R})$: The set of all $m \times n$ matrices with real entries forms a vector space. Matrix addition and scalar multiplication satisfy all vector space axioms. This example generalizes $\mathbb{R}^n$ and plays a key role in linear transformations.

The fact that such diverse mathematical objects, from simple number lists to entire functions and polynomials, can all be treated as vectors is a powerful idea. By abstracting the notion of "vector," we gain a unifying language that allows us to transfer intuition and techniques across different areas of mathematics. In later sections, we will see how this framework lets us study polynomials and functions using the same ideas we developed for arrows in space.

3 Inner Product Spaces

Motivation from the Dot Product

In Euclidean space $\mathbb{R}^n$, one of the most fundamental operations between two vectors is the **dot product**. Given two vectors $\vec{v}, \vec{w} \in \mathbb{R}^n$, the dot product is defined as:

$$\vec{v} \cdot \vec{w} = \sum_{i=1}^{n} v_i w_i$$

This operation is algebraically convenient, but it also has deep geometric meaning. In particular, the dot product relates to the angle θ between two vectors:

$$\vec{v} \cdot \vec{w} = \|\vec{v}\|\|\vec{w}\| \cos\theta$$

From this formula, we immediately get an important condition for **orthogonality**: two vectors are **orthogonal** (or perpendicular) if and only if their dot product is zero:

$$\vec{v} \cdot \vec{w} = 0 \quad \Leftrightarrow \quad \theta = \frac{\pi}{2}$$

This concept of orthogonality is central to many parts of linear algebra and functional analysis. But it turns out the standard definition of a *vector space* does not include any notion of length or angle. There is no multiplication between vectors in the vector space axioms.

So to capture notions like length, angle, and orthogonality, we must define a new kind of structure, an **inner product space**.

Definition: Inner Product Space

An **inner product space** is a vector space equipped with an additional operation that acts like the dot product.

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

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

satisfying the following axioms for all $\vec{u}, \vec{v}, \vec{w} \in V$ and scalars $a \in \mathbb{F}$:

1. **Conjugate symmetry:** $\langle \vec{v}, \vec{u} \rangle = \overline{\langle \vec{u}, \vec{v} \rangle}$
2. **Linearity in the first slot:** $\langle a\vec{u} + \vec{v}, \vec{w} \rangle = a\langle \vec{u}, \vec{w} \rangle + \langle \vec{v}, \vec{w} \rangle$
3. **Positive-definiteness:** $\langle \vec{v}, \vec{v} \rangle \geq 0$, and $\langle \vec{v}, \vec{v} \rangle = 0 \iff \vec{v} = \vec{0}$

When a vector space is equipped with such an inner product, it is called an **inner product space**.

Examples of Inner Product Spaces

- **Euclidean Space** $\mathbb{R}^n$: The standard inner product is the dot product:

$$\langle \vec{v}, \vec{w} \rangle = \sum_{i=1}^{n} v_i w_i$$

  This is the familiar case taught in basic linear algebra courses.

- **Complex Space** $\mathbb{C}^n$: The inner product is defined by:

$$\langle \vec{v}, \vec{w} \rangle = \sum_{i=1}^{n} v_i \overline{w_i}$$

  The complex conjugate ensures that $\langle \vec{v}, \vec{v} \rangle$ is always a non-negative real number.

- **Function Space** $C[a, b]$: The set of continuous real-valued functions on an interval $[a, b]$ can be turned into an inner product space by defining:

$$\langle f, g \rangle = \int_a^b f(x)g(x)\, dx$$

  This is often called the L^2 inner product. It allows us to measure the "overlap" between functions, just like the dot product measures alignment between vectors.

- **Polynomial Space** $\mathbb{P}_n$: The set of real polynomials of degree at most n can be given an inner product using an integral over an interval:

$$\langle p, q \rangle = \int_a^b p(x)q(x)w(x)\, dx$$

  where $w(x)$ is a weight function. This setting leads directly to the study of orthogonal polynomials.

Length and Orthogonality in Inner Product Spaces

Just like in Euclidean space, the inner product allows us to define:

- **Length (norm):**

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

- **Orthogonality:** Two vectors $\vec{v}, \vec{w}$ are orthogonal if:

$$\langle \vec{v}, \vec{w} \rangle = 0$$

These concepts extend naturally to spaces of functions and polynomials. This is what allows us to speak of *orthogonal polynomials*, as we'll explore in the next section.

Why This Matters

By abstracting the dot product into a general inner product, we now have a language to talk about angle, projection, and perpendicularity even in infinite-dimensional settings. This framework is the foundation of Fourier analysis, quantum mechanics, and many areas of applied mathematics.

Most importantly for this paper, inner product spaces allow us to define when two functions or polynomials are orthogonal, unlocking the door to approximation, spectral theory, and beautiful structure in classical mathematical physics.

4 Defining Orthogonal Polynomials

We now have all the mathematical machinery we need to define one of the central concepts of this paper: **orthogonal polynomials**.

What Are Orthogonal Polynomials?

Just as vectors can be orthogonal if their inner product is zero, so too can polynomials. We define two polynomials $p(x)$ and $q(x)$ to be orthogonal with respect to a weight function $w(x)$ over an interval $[a, b]$ if:

$$\langle p, q \rangle = \int_a^b p(x)q(x)w(x)\,dx = 0$$

A **sequence of orthogonal polynomials** $\{p_0(x), p_1(x), p_2(x), \dots\}$ is a sequence where each $p_n(x)$ is a polynomial of degree n, and the set satisfies:

$$\langle p_n, p_m \rangle = 0 \quad \text{for } n \neq m$$

The weight function $w(x)$ must be positive on the interval $[a, b]$, and it can influence the form of the polynomials significantly. The choice of interval and weight function determines the inner product, and thus determines what "orthogonal" means in that space.

A Simple Example of Orthogonality

Let's consider two simple polynomials: $p(x) = 1$ and $q(x) = x$. Are they orthogonal with respect to the standard inner product on $[-1, 1]$, using the weight function $w(x) = 1$?

We compute:

$$\langle 1, x \rangle = \int_{-1}^{1} (1)(x)(1)\,dx = \int_{-1}^{1} x\,dx = 0$$

So yes, the constant function 1 and the linear function x are orthogonal under this inner product. This example highlights how even familiar functions can be orthogonal in this extended sense.

Building Orthogonal Polynomials: Regular Method

Now, instead of using special techniques (like Gram-Schmidt), we can construct orthogonal polynomials directly by choosing a specific inner product space, defined by an interval and a weight function, and checking orthogonality.

Example 1: Orthogonal polynomials on $[-1, 1]$ with weight $w(x) = 1$

Let's check if $p_0(x) = 1$, $p_1(x) = x$, and $p_2(x) = \frac{1}{2}(3x^2 - 1)$ are mutually orthogonal.

We already showed that:

$$\langle p_0, p_1 \rangle = \int_{-1}^{1} 1 \cdot x\,dx = 0$$

Now:

$$\langle p_0, p_2 \rangle = \int_{-1}^{1} 1 \cdot \frac{1}{2}(3x^2 - 1)\, dx = \frac{1}{2} \int_{-1}^{1} (3x^2 - 1)\, dx = \frac{1}{2} \left(3 \int_{-1}^{1} x^2 dx - \int_{-1}^{1} dx \right)$$

$$= \frac{1}{2} \left(3 \cdot \frac{2}{3} - 2 \right) = \frac{1}{2}(2 - 2) = 0$$

Also:

$$\langle p_1, p_2 \rangle = \int_{-1}^{1} x \cdot \frac{1}{2}(3x^2 - 1)\, dx = \frac{1}{2} \int_{-1}^{1} x(3x^2 - 1)\, dx = \frac{1}{2} \int_{-1}^{1} (3x^3 - x)\, dx = 0$$

So, these three polynomials are mutually orthogonal over $[-1, 1]$ with weight 1. Even though we didn't derive them algorithmically, they satisfy the required condition.

Example 2: Orthogonal polynomials on $[0, 1]$ with weight $w(x) = x$

Let's test:

$$p_0(x) = 1, \quad p_1(x) = x, \quad p_2(x) = x^2 - \frac{6}{5}$$

We check:

$$\langle p_0, p_1 \rangle = \int_{0}^{1} 1 \cdot x \cdot x\, dx = \int_{0}^{1} x^2 dx = \frac{1}{3} \quad \Rightarrow \quad \text{Not orthogonal}$$

So we adjust $p_1(x) = x - \frac{3}{4}$, and recompute:

$$\langle p_0, p_1 \rangle = \int_{0}^{1} 1 \cdot (x - \frac{3}{4}) \cdot x\, dx = \int_{0}^{1} x(x - \frac{3}{4}) dx = \int_{0}^{1} (x^2 - \frac{3}{4}x) dx = \frac{1}{3} - \frac{3}{8} = 0$$

So, $p_0(x) = 1$, $p_1(x) = x - \frac{3}{4}$ are orthogonal with respect to the inner product:

$$\langle f, g \rangle = \int_{0}^{1} f(x) g(x) x\, dx$$

This shows how changing the weight function modifies what it means to be orthogonal.

Some Observations

- Orthogonal polynomials are highly sensitive to the choice of interval and weight function.
- Even "familiar" polynomials like x^2 or $x - \frac{1}{2}$ may or may not be orthogonal depending on the inner product.
- The orthogonality condition is symmetric: if $\langle p, q \rangle = 0$, then $\langle q, p \rangle = 0$.
- This construction is foundational for approximation theory, as orthogonal polynomials act like "basis vectors" in function spaces.

Why This Matters

We now see that polynomials can behave like geometric vectors. They can be "perpendicular," "normalized," and combined just like arrows in space. Orthogonality gives us structure in the infinite world of functions and polynomials.

In later sections, we will explore how to systematically generate orthogonal polynomials using algebraic methods like the Gram-Schmidt process, and we will study their deeper properties including recurrence relations and root behavior.

5 Gram-Schmidt Process

The Gram-Schmidt process is a fundamental algorithm in linear algebra that allows us to construct an **orthogonal basis** from any given basis in an inner product space. While a given basis spans the space, the Gram-Schmidt process transforms it into a more "structured" basis, one where all the vectors are mutually orthogonal.

Intuition Behind the Process

At its core, the process works by taking each new vector and subtracting off the projections of that vector onto the ones that came before it, leaving only the "new direction" that wasn't already captured. It is a method of peeling off components until what remains is orthogonal to everything that came earlier.

Example: Vectors in $\mathbb{R}^3$

Let's orthogonalize the following basis of $\mathbb{R}^3$:

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

We begin with $\vec{u}_1 = \vec{v}_1$.

Next, we subtract the projection of $\vec{v}_2$ onto $\vec{u}_1$:

$$\vec{u}_2 = \vec{v}_2 - \frac{\langle \vec{v}_2, \vec{u}_1 \rangle}{\langle \vec{u}_1, \vec{u}_1 \rangle} \vec{u}_1$$

Compute:

$$\langle \vec{v}_2, \vec{u}_1 \rangle = 1(1) + 0(1) + 1(0) = 1, \quad \langle \vec{u}_1, \vec{u}_1 \rangle = 1^2 + 1^2 + 0^2 = 2$$

$$\vec{u}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix}$$

Repeat for $\vec{u}_3$, subtracting projections onto both $\vec{u}_1$ and $\vec{u}_2$.

This results in an orthogonal basis $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$. If desired, each can be normalized to form an orthonormal basis.

Applying Gram-Schmidt to Polynomials

Since we've already seen that polynomials can form a vector space and can be equipped with an inner product (such as integration over an interval), we can apply the same Gram-Schmidt process to construct orthogonal polynomials.

Let's start with the standard basis for the space of polynomials of degree at most 2:

$$\{1, x, x^2\}$$

We'll construct orthogonal polynomials over the interval $[-1, 1]$, with weight function $w(x) = 1$. The inner product is:

$$\langle f, g \rangle = \int_{-1}^{1} f(x) g(x) \, dx$$

Example: Constructing Orthogonal Polynomials

Let:

$$p_0(x) = 1$$

$$p_1(x) = x - \frac{\langle x, p_0 \rangle}{\langle p_0, p_0 \rangle} p_0 = x - \frac{\int_{-1}^{1} x \cdot 1 \, dx}{\int_{-1}^{1} 1^2 \, dx} \cdot 1 = x - \frac{0}{2} = x$$

So far:

$$\langle p_0, p_1 \rangle = \int_{-1}^{1} 1 \cdot x \, dx = 0$$

Now define:

$$p_2(x) = x^2 - \frac{\langle x^2, p_0 \rangle}{\langle p_0, p_0 \rangle} p_0 - \frac{\langle x^2, p_1 \rangle}{\langle p_1, p_1 \rangle} p_1$$

Compute:

$$\langle x^2, p_0 \rangle = \int_{-1}^{1} x^2 \cdot 1 \, dx = \frac{2}{3}, \quad \langle p_0, p_0 \rangle = 2 \Rightarrow \text{proj}_{p_0}(x^2) = \frac{2}{3}/2 = \frac{1}{3}$$

$$\langle x^2, p_1 \rangle = \int_{-1}^{1} x^3 \, dx = 0, \quad \langle p_1, p_1 \rangle = \int_{-1}^{1} x^2 dx = \frac{2}{3} \Rightarrow \text{proj}_{p_1}(x^2) = 0$$

So:

$$p_2(x) = x^2 - \frac{1}{3}$$

Final orthogonal polynomials:

$$p_0(x) = 1, \quad p_1(x) = x, \quad p_2(x) = x^2 - \frac{1}{3}$$

These are mutually orthogonal under the given inner product.

Another Example: Weight Function $w(x) = 1 + x$ on $[0, 1]$

Let's try constructing the first two orthogonal polynomials with a different weight function:

$$\langle f, g \rangle = \int_{0}^{1} f(x) g(x) (1 + x) \, dx$$

Let $p_0(x) = 1$
Now compute:

$$p_1(x) = x - \frac{\langle x, 1 \rangle}{\langle 1, 1 \rangle}$$

$$\langle x, 1 \rangle = \int_{0}^{1} x(1 + x) dx = \int_{0}^{1} (x + x^2) dx = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}$$

$$\langle 1, 1 \rangle = \int_{0}^{1} (1 + x) dx = 1 + \frac{1}{2} = \frac{3}{2}$$

So:

$$p_1(x) = x - \frac{5}{6} \cdot \frac{2}{3} = x - \frac{10}{18} = x - \frac{5}{9}$$

Even this simple step shows how the weight function completely alters the result. Gram-Schmidt for polynomials is exact but can become algebraically messy.

6 Examples of Orthogonal Polynomials

Orthogonal polynomials appear throughout pure and applied mathematics, particularly in approximation theory, numerical analysis, and mathematical physics. In this section, we explore specific families of orthogonal polynomials and their properties, beginning with the Legendre polynomials.

Legendre Polynomials

The **Legendre polynomials** $\{P_n(x)\}$ form a sequence of orthogonal polynomials on the interval $[-1, 1]$ with weight function $w(x) = 1$. That is:

$$\int_{-1}^{1} P_n(x)P_m(x)\,dx = 0 \quad \text{for } n \neq m$$

They arise naturally when solving certain differential equations in physics, especially problems exhibiting spherical symmetry.

First Few Polynomials

We list the first few Legendre polynomials:

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= x \\ P_2(x) &= \frac{1}{2}(3x^2 - 1) \\ P_3(x) &= \frac{1}{2}(5x^3 - 3x) \\ P_4(x) &= \frac{1}{8}(35x^4 - 30x^2 + 3) \end{aligned}$$

Each $P_n(x)$ is a degree-n polynomial, and they are orthogonal under the standard inner product on $[-1, 1]$.

Orthogonality Property

The orthogonality relation is:

$$\int_{-1}^{1} P_n(x)P_m(x)\,dx = \begin{cases} 0 & \text{if } n \neq m \\ \frac{2}{2n+1} & \text{if } n = m \end{cases}$$

This property allows Legendre polynomials to serve as a basis for approximating functions defined on the interval $[-1, 1]$, similar to how Fourier series are built from sine and cosine functions.

Appearance in Physics: Solving Laplace's Equation

One of the most significant appearances of Legendre polynomials is in solving **Laplace's equation** in spherical coordinates:

$$\nabla^2\Phi = 0$$

In spherical coordinates (r, θ, ϕ), assuming azimuthal symmetry (no ϕ-dependence), Laplace's equation reduces to:

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\Phi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Phi}{\partial\theta}\right) = 0$$

Using separation of variables $\Phi(r, \theta) = R(r)\Theta(\theta)$, the angular part satisfies:

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right)+\lambda\Theta=0$$

Letting $x=\cos\theta$, this becomes:

$$\frac{d}{dx}\left[(1-x^2)\frac{d\Theta}{dx}\right]+\lambda\Theta=0$$

This is precisely the **Legendre differential equation**:

$$\frac{d}{dx}\left[(1-x^2)\frac{dP}{dx}\right]+\lambda P=0$$

This equation has solutions only for $\lambda=n(n+1)$, where $P(x)=P_n(x)$ is the Legendre polynomial of degree n.

Legendre's Differential Equation

To summarize, Legendre polynomials are solutions to the second-order linear differential equation:

$$(1-x^2)\frac{d^2P_n}{dx^2}-2x\frac{dP_n}{dx}+n(n+1)P_n=0$$

This equation arises both from mathematical properties of orthogonality and from physical systems involving radial and angular symmetry — most notably when solving for the potential field around spherical objects.

Chebyshev Polynomials

The **Chebyshev polynomials of the first kind**, denoted $\{T_n(x)\}$, are orthogonal polynomials defined on the interval $[-1,1]$ with respect to the weight function:

$$w(x)=\frac{1}{\sqrt{1-x^2}}$$

That is, they satisfy the orthogonality condition:

$$\int_{-1}^{1}\frac{T_n(x)T_m(x)}{\sqrt{1-x^2}}\,dx=0\quad\text{for } n\neq m$$

First Few Polynomials

The first few Chebyshev polynomials of the first kind are:

$$\begin{aligned}
T_0(x)&=1\\
T_1(x)&=x\\
T_2(x)&=2x^2-1\\
T_3(x)&=4x^3-3x\\
T_4(x)&=8x^4-8x^2+1
\end{aligned}$$

Each $T_n(x)$ is a degree-n polynomial, and these polynomials are highly symmetric and computationally efficient for certain numerical tasks.

Trigonometric Definition

One of the most striking features of Chebyshev polynomials is their connection to trigonometry. They can be defined using the identity:

$$T_n(x) = \cos(n \cos^{-1}(x)), \quad x \in [-1, 1]$$

This definition reveals their oscillatory behavior and helps explain why they are especially effective in approximating functions. Since cosine is bounded between -1 and 1, the values of $T_n(x)$ also remain bounded on the interval $[-1, 1]$, making them numerically stable and well-suited for interpolation and minimax approximation.

Importance in Approximation Theory

Chebyshev polynomials are crucial in numerical analysis due to their minimization of the maximum error in polynomial interpolation, a problem known as the **minimax approximation**. If we want to approximate a function $f(x)$ with a polynomial, choosing the interpolation nodes to be the zeros of Chebyshev polynomials significantly reduces the error due to Runge's phenomenon.

Hermite Polynomials

The **Hermite polynomials**, denoted $\{H_n(x)\}$, are orthogonal polynomials on the interval $(-\infty, \infty)$ with respect to the weight function:

$$w(x) = e^{-x^2}$$

They satisfy the orthogonality condition:

$$\int_{-\infty}^{\infty} H_n(x) H_m(x) e^{-x^2}\, dx = 0 \quad \text{for } n \neq m$$

First Few Hermite Polynomials

The first few Hermite polynomials are:

$$\begin{aligned} H_0(x) &= 1 \\ H_1(x) &= 2x \\ H_2(x) &= 4x^2 - 2 \\ H_3(x) &= 8x^3 - 12x \\ H_4(x) &= 16x^4 - 48x^2 + 12 \end{aligned}$$

These polynomials are symmetric in structure and increase rapidly in degree and coefficient size.

Hermite Differential Equation

Hermite polynomials are the solutions to the second-order linear differential equation:

$$H_n''(x) - 2xH_n'(x) + 2nH_n(x) = 0$$

This is known as the **Hermite differential equation**. Its structure, involving both the function and its first and second derivatives, makes it well-suited to describing oscillatory behavior with damping, a key reason why it emerges in quantum mechanics.

Appearance in Quantum Mechanics: The 1D Harmonic Oscillator

In quantum mechanics, the time-independent Schrödinger equation for the one-dimensional harmonic oscillator is:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \frac{1}{2}m\omega^2 x^2\psi = E\psi$$

By applying a change of variables and using dimensionless units, this equation can be rewritten as:

$$\frac{d^2\psi}{dy^2} = (y^2 - \lambda)\psi$$

To solve this, physicists typically look for solutions of the form:

$$\psi_n(x) = N_n \cdot H_n(x) \cdot e^{-x^2/2}$$

Here:
$H_n(x)$ is the Hermite polynomial of degree n - $e^{-x^2/2}$

Laguerre Polynomials

The **Laguerre polynomials**, denoted $\{L_n(x)\}$, are orthogonal polynomials on the interval $[0, \infty)$ with respect to the weight function:

$$w(x) = e^{-x}$$

They satisfy the orthogonality relation:

$$\int_0^\infty L_n(x)L_m(x)e^{-x}\,dx = 0 \quad \text{for } n \neq m$$

First Few Laguerre Polynomials

The first few Laguerre polynomials are:

$$\begin{aligned}
L_0(x) &= 1 \\
L_1(x) &= -x + 1 \\
L_2(x) &= \frac{1}{2}(x^2 - 4x + 2) \\
L_3(x) &= \frac{1}{6}(-x^3 + 9x^2 - 18x + 6)
\end{aligned}$$

Each $L_n(x)$ is a polynomial of degree n, and the signs alternate in a consistent pattern. These polynomials are often used in applications involving exponential decay.

Laguerre Differential Equation

Laguerre polynomials satisfy the differential equation:

$$x\frac{d^2L_n}{dx^2} + (1 - x)\frac{dL_n}{dx} + nL_n = 0$$

This equation characterizes $L_n(x)$ as solutions that remain finite as $x \to \infty$, making them ideal for systems with boundary conditions on $[0, \infty)$.

Appearance in Quantum Mechanics: Hydrogen Atom

Laguerre polynomials appear naturally in quantum mechanics when solving the Schrödinger equation for the hydrogen atom, particularly in the radial component.

The time-independent Schrödinger equation in spherical coordinates for the hydrogen atom separates into radial and angular parts. The radial part becomes:

$$\frac{d^2u}{dr^2} + \left[\frac{2m}{\hbar^2}\left(E + \frac{e^2}{4\pi\epsilon_0 r}\right) - \frac{\ell(\ell+1)}{r^2}\right]u = 0$$

After changing variables and applying appropriate substitutions, this reduces to a form whose solutions involve a polynomial term that turns out to be a generalized Laguerre polynomial.

The radial wavefunction $R_{n\ell}(r)$ has the form:

$$R_{n\ell}(r) = N_{n\ell} \cdot r^\ell \cdot e^{-r/na_0} \cdot L_{n-\ell-1}^{2\ell+1}\left(\frac{2r}{na_0}\right)$$

Here: $L_k^{(\alpha)}(x)$ is a generalized Laguerre polynomial , $N_{n\ell}$ is a normalization constant , a_0 is the Bohr radius

7 Approximating Functions

One of the most powerful uses of orthogonal polynomials is in approximating arbitrary functions. Just as we use sine and cosine functions in Fourier series, we can use orthogonal polynomials as a basis to approximate a function over a given interval.

Orthogonal Polynomial Approximation

Given a function $f(x)$ defined on an interval $[a, b]$, and a set of orthogonal polynomials $\{p_n(x)\}$ with respect to a weight function $w(x)$, we can approximate f as:

$$f(x) \approx \sum_{n=0}^{N} c_n p_n(x)$$

where the coefficients c_n are given by:

$$c_n = \frac{\langle f, p_n\rangle}{\langle p_n, p_n\rangle} = \frac{\int_a^b f(x)p_n(x)w(x)\,dx}{\int_a^b p_n^2(x)w(x)\,dx}$$

This is analogous to projecting a vector onto a basis, except here, our "vectors" are functions, and our inner product is defined using an integral.

Comparison to Taylor Series

The Taylor series of a function at a point $x = a$ provides a polynomial approximation using derivatives evaluated at a single point:

$$f(x) \approx \sum_{n=0}^{N} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

While Taylor series are highly effective near the expansion point, they can perform poorly over wider intervals, especially if the function has rapid changes or lacks smoothness.

Orthogonal polynomial approximation, on the other hand, minimizes the approximation error over an entire interval. This makes it much more effective when global behavior of the function is important.

Chebyshev Approximation and Least Squares Error

Chebyshev polynomials are particularly important because they minimize the **maximum deviation** (also known as minimax error) from the function being approximated. In other words, the best polynomial approximation $P_n(x)$ to a function $f(x)$ under the maximum norm satisfies:

$$\|f - P_n\|_\infty = \min_{Q \in \mathcal{P}_n} \|f - Q\|_\infty$$

where $\mathcal{P}_n$ is the set of all polynomials of degree $\leq n$, and the infinity norm represents the maximum error over the interval.

The Chebyshev nodes (the roots of the Chebyshev polynomials) are used to choose interpolation points that minimize Runge's phenomenon and yield near-optimal approximation:

$$x_k = \cos\left(\frac{2k-1}{2n}\pi\right), \quad k = 1, 2, ..., n$$

Example: Approximating $|x|$ using Chebyshev Polynomials

Let us consider the function $f(x) = |x|$ defined on the interval $[-1, 1]$. This function is continuous, but not differentiable at $x = 0$, which makes it poorly approximated by a Taylor series centered at the origin.

To approximate f using Chebyshev polynomials $T_n(x)$, we compute the coefficients:

$$c_n = \frac{2}{\pi}\int_{-1}^{1} \frac{f(x)T_n(x)}{\sqrt{1-x^2}}\,dx, \quad \text{for } n \geq 1, \quad c_0 = \frac{1}{\pi}\int_{-1}^{1} \frac{f(x)}{\sqrt{1-x^2}}\,dx$$

For $f(x) = |x|$, the first few Chebyshev coefficients are known to be:

$$\begin{aligned} c_0 &= \frac{2}{\pi} \\ c_1 &= 0 \\ c_2 &= -\frac{4}{3\pi} \\ c_3 &= 0 \\ c_4 &= \frac{4}{15\pi} \end{aligned}$$

Thus, the degree-4 Chebyshev approximation of $|x|$ is:

$$f(x) \approx \frac{2}{\pi}T_0(x) - \frac{4}{3\pi}T_2(x) + \frac{4}{15\pi}T_4(x)$$

Substituting the Chebyshev polynomials:

$$\begin{aligned} T_0(x) &= 1 \\ T_2(x) &= 2x^2 - 1 \\ T_4(x) &= 8x^4 - 8x^2 + 1 \end{aligned}$$

We obtain:

$$f(x) \approx \frac{2}{\pi} - \frac{4}{3\pi}(2x^2 - 1) + \frac{4}{15\pi}(8x^4 - 8x^2 + 1)$$

Why This Works

This approximation captures the shape of $|x|$ over the entire interval, even though $|x|$ is not smooth at $x = 0$. Unlike the Taylor series, which relies on derivatives at a point and fails to represent the global shape accurately, Chebyshev approximation spreads the approximation power across the whole interval $[-1, 1]$.

In the figure, we compare the Chebyshev approximation and the Taylor approximation of $|x|$ up to degree 5. The Taylor polynomial oscillates and diverges near the endpoints, while the Chebyshev approximation remains close to $|x|$ over the whole interval.

Mean Square Error (L2 Norm)

Another common measure of how well a polynomial approximates a function is the **mean square error**, defined as:

$$E = \int_a^b (f(x) - P_n(x))^2 w(x)\, dx$$

Orthogonal polynomials minimize this error when projected using the inner product structure. That is, choosing $P_n(x) = \sum_{k=0}^{n} c_k p_k(x)$ with coefficients derived via orthogonal projection yields the polynomial of best approximation in the L^2 norm.

8 Recurrence Relations and Derivative Expressions

Orthogonal polynomials are not only defined by their orthogonality conditions and weight functions, but also by elegant recursive and differential structures. These recurrence relations and derivative identities allow us to construct polynomials efficiently and understand their properties in depth.

Recurrence Relations

Many classical orthogonal polynomials satisfy three-term recurrence relations of the form:

$$p_{n+1}(x) = (A_n x + B_n) p_n(x) + C_n p_{n-1}(x)$$

These allow each polynomial to be built using only the previous two.

Legendre Polynomials:

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x)$$

Chebyshev Polynomials (First Kind):

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

Hermite Polynomials:

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x)$$

Laguerre Polynomials:

$$(n+1)L_{n+1}(x) = (2n+1-x)L_n(x) - nL_{n-1}(x)$$

Rodrigues' Theorem

Rodrigues' formula provides an explicit representation for orthogonal polynomials in terms of derivatives:

$$p_n(x) = \frac{1}{w(x)} \cdot \frac{d^n}{dx^n} \left[w(x) \cdot r(x)^n \right]$$

where $w(x)$ is the weight function and $r(x)$ is a polynomial chosen based on the family.

Examples:

- **Legendre:**

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n$$

- **Hermite:**

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} \left(e^{-x^2} \right)$$

- **Laguerre:**
$$L_n(x) = \frac{e^x}{n!}\frac{d^n}{dx^n}(x^n e^{-x})$$
- **Chebyshev:** While Chebyshev polynomials do not have a traditional Rodrigues formula, they can be defined via:
$$T_n(x) = \cos(n \cos^{-1}(x))$$

9 Roots

Orthogonal polynomials possess highly structured roots. These roots play crucial roles in approximation theory, numerical integration, and interpolation.

Real, Simple Roots within the Interval

Let $\{p_n(x)\}$ be a sequence of polynomials orthogonal on an interval $[a, b]$ with respect to a positive weight function $w(x) > 0$. Then:

- $p_n(x)$ has exactly n real and distinct roots.
- All roots lie strictly within the open interval (a, b).

Sketch of Proof: Assume $p_n(x)$ has fewer than n distinct real roots in (a, b). Then there exists a nonzero polynomial $q(x)$ of degree $< n$ that shares the same sign as $p_n(x)$ on $[a, b]$, meaning:

$$\int_a^b p_n(x)q(x)w(x)\,dx \neq 0$$

This contradicts orthogonality, which requires $\langle p_n, q\rangle = 0$ for all $\deg(q) < n$. Thus, p_n must have exactly n distinct real roots in (a, b).

Interlacing of Roots

Another remarkable property is that the roots of consecutive orthogonal polynomials interlace. That is, between any two roots of $p_n(x)$, there exists exactly one root of $p_{n+1}(x)$.

Theorem: Let $p_n(x)$ and $p_{n+1}(x)$ be orthogonal polynomials with respect to the same weight function on $[a, b]$. Then the roots of $p_n(x)$ and $p_{n+1}(x)$ interlace.

Proof: Let $x_1 < x_2 < \cdots < x_n$ be the n distinct roots of $p_n(x)$. Define the function:

$$f(x) = \frac{p_{n+1}(x)}{p_n(x)}$$

Since p_n has simple real roots, $f(x)$ is continuous and strictly monotonic on each subinterval between consecutive roots of $p_n(x)$. It follows from Sturm comparison or intermediate value arguments that $f(x)$ must change sign between each x_i, implying $p_{n+1}(x) = 0$ has exactly one root between each pair x_i, x_{i+1}.

Thus, $p_{n+1}(x)$ has n roots between the roots of $p_n(x)$, and one additional root outside the interval spanned by the roots of $p_n(x)$, ensuring the full interlacing structure:

$$x_1^{(n+1)} < x_1^{(n)} < x_2^{(n+1)} < x_2^{(n)} < \cdots < x_n^{(n)} < x_{n+1}^{(n+1)}$$

References

[1] Francisco Marcellán and Yuan Xu. *On the theory of orthogonal polynomials*. arXiv preprint arXiv:1303.2825, 2013. https://arxiv.org/pdf/1303.2825

[2] M. T. J. Johnson. *Orthogonal Polynomials Overview*. University of Sydney. https://www.sydney.edu.au/science/chemistry/ mjtj/CHEM3117/Resources/poly$_e$*tc.pdf*

[3] Zhengfu Xu. *Orthogonal Polynomials Lecture Notes*. University of Notre Dame. https://www3.nd.edu/ zxu2/acms40390F13/Lec-8.2.pdf

[4] Abby Ward. *Orthogonal Polynomials Summary*. University of Chicago. https://people.cs.uchicago.edu/ laci/HANDOUTS/abbyward-orthog.pdf

[5] Xiaoming Jia. *Orthogonal Polynomials Notes*. Iowa State University. https://faculty.sites.iastate.edu/jia/files/inline-files/orthogonal-polys.pdf

[6] Bei Wang. *Numerical PDEs – Lecture 20: Orthogonal Polynomials*. University of Utah. https://www.sci.utah.edu/ beiwang/teaching/cs6210-fall-2016/lecture20.pdf