Orthogonal Polynomials and Polynomial Approximation to Functions

1 The Vector Space $L_2(a, b)$

The vector space $L_2(a, b)$ is the set of all square-integrable functions defined over the interval $[a, b]$. If a function $f \in L_2(a, b)$, then

$$\int_a^b dx\ f(x) < \infty \tag{1}$$

Given two such function f and g, it is easy to see that a linear combination $c_1 f + c_2 g$ also satisfies condition (1). We can define a 'zero' function f_0 such that $f_0(x) = 0\ \forall\ x \in [a, b]$. Then, it is easy to see that the set of all such functions form a vector space. We can further define an inner-product on this space as follows

$$f \cdot g = \int_a^b dx\ f(x) g(x) \tag{2}$$

It is easy to check that this satisfies the linearity condition for the inner product

$$f \cdot (\alpha g + \beta h) = \alpha f \cdot g + \beta f \cdot h \tag{3}$$

where α, β are numbers and f, g, h are functions in $L_2(a, b)$. This inner product also induces a norm ('length' of vector f) as follows

$$\begin{aligned} ||f||^2 &= f \cdot f \qquad (4) \\ &= \int_a^b dx\ f^2(x) \end{aligned}$$

Since there is an infinite number of linearly independent functions which satisfy condition (1), this is an infinite dimensional vector space.

2 The Vector Space $\mathcal{P}_n$

We now consider the set of all polynomials of degree less than equal to n over the interval $[-1, 1]$. We will see how to generalize the analysis to an arbitrary interval $[a, b]$. It is easy to see that the set of all such polynomials also forms a vector space, which we denote as $\mathcal{P}_n$. This is clearly a sub-space of the vector space $L_2(-1, 1)$. It is also easy to see that this is an $n + 1$ dimensional vector space, since the set $\{1, x, x^2, x^3, ..., x^n\}$, which is linearly independent, spans the space (any such polynomial can be written as a linear combination of these functions) and therefore forms a basis, since any such polynomial can be expressed as

$$p_n(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + ... + c_n x^n \tag{5}$$

However, this basis is not orthonormal. We can construct an orthonormal basis out of this basis using Gram-Schmidt orthonormalization. Let us call the basis $\{\phi_0 = 1, \phi_1 = x, \phi_2 = x^2, .., \phi_n = x^n\}$ the 'natural' basis, since we naturally view a polynomial of degree n as an expansion in these simple functions.

Let us denote the orthonormal basis as $\{P_0, P_1, .., P_n\}$. Following the Gram-Schmidt orthonormalization procedure, we choose the first unit 'vector' as ϕ_0 itself, normalized

$$\begin{aligned} P_0 &= \frac{1}{||\phi_0||}\ \phi_0 \\ &= \frac{1}{\sqrt{2}} \end{aligned}$$

where

$$\begin{aligned} ||\phi_0||^2 &= \phi_0 \cdot \phi_0 \\ &= \int_{-1} 1dx\ 1 \\ &= 2 \end{aligned}$$

Next, we add the vector ϕ_1 and project out its component along the unit-vector P_0. Let us call this new vector ψ_1

$$\begin{aligned} \psi_1 &= \phi_1 - (\phi_1 \cdot P_0)\ P_0 \\ &= \phi_1 \end{aligned}$$

since $\phi_1 \cdot P_0 = 0$. Then, the second unit vector is just

$$\begin{aligned} P_1 &= \frac{1}{||\psi_1||}\ \psi_1 \\ &= \sqrt{\frac{3}{2}}\ x \end{aligned}$$

Next, we take ϕ_2 and project out its components along P_0 and P_1 to give ψ_2

$$\begin{aligned} \psi_2 &= \phi_2 - (\phi_2 \cdot P_1)\ P_1 - (\phi_2 \cdot P_0)\ P_0 \\ &= \phi_2 - (\phi_2 \cdot P_0)\ P_0 \\ &= x^2 - \frac{1}{3} \end{aligned}$$

Then, we get the unit vector P_2 as

$$\begin{aligned} P_2 &= \frac{1}{||\psi_2||}\ \psi_2 \\ &= \sqrt{\frac{5}{8}}\ (3x^2 - 1) \end{aligned}$$

We can similarly recursively generate the unit vectors $P_3, P_4, .., P_n$. These orthonormal functions defined over the interval $[-1, 1]$ are just the 'special' functions known as *Legendre Polynomials*. The next two orthonormal vectors in this set are

$$\begin{aligned} P_3 &= \frac{1}{2}\sqrt{\frac{7}{2}}\ (5x^3 - 3x) \\ P_4 &= \frac{1}{8}\sqrt{\frac{9}{2}}\ (35x^4 - 30x^2 + 3) \end{aligned}$$

3 Polynomial Approximation

Consider an arbitrary function f belonging to $L_2(-1, 1)$. Say, we wish to approximate this function by a polynomial of degree n. To do this, of all the polynomials of degree less than equal to n, we wish to pick the one which is 'closest' to the function f. Since each such polynomial belongs to the subspace $\mathcal{P}_n$ of the vector space $L_2(-1,1)$, the natural measure of 'closeness' is the norm $||f - p_n||$ where $p_n \in \mathcal{P}_n$, which can be thought of as the 'distance' between functions f and p_n. The mathematical problem then reduces to determining $p_n \in \mathcal{P}_n$ which minimises this norm. Since $\mathcal{P}_n$ is a subspace of $L_2(-1,1)$. the element of $\mathcal{P}_n$ which minimises this norm is just the orthogonal projection of the vector f on to $\mathcal{P}_n$. This is just the function/vector $f_\perp$ given by

$$f_\perp = \sum_{i=0}^{n} (f \cdot P_i)\ P_i$$

where

$$f \cdot P_i = \int_{-1}^{1} dx\ f(x) P_i(x) \tag{6}$$

The strategy can be extended to an arbitrary interval $x \in [a,b]$. First, we define a new variable y as follows

$$y = \left(\frac{2}{b-a}\right) x + \left(\frac{a+b}{a-b}\right) \tag{7}$$

This maps y to the interval $[-1,1]$. Next, we define the function

$$g(y) = f\left[\left(\frac{b-a}{2}\right) y + \left(\frac{a+b}{2}\right)\right] \tag{8}$$

This is defined over $[-1,1]$. We can approximate $g(y)$ by its orthogonal projection over $\mathcal{P}_n$, $g_\perp$. Then, the required polynomial approximation for f is

$$f_\perp = g_\perp\left[\left(\frac{2}{b-a}\right) x + \left(\frac{a+b}{a-b}\right)\right] \tag{9}$$

4 Generalizing the Inner Product: The Weight Function

The definition (2) of the inner product can be generalized as follows: Given a non-negative function $w(x)$ over the interval $[a,b]$, we can generalize the definition of inner product to

$$f \cdot g = \int_{a}^{b} dx\ w(x) f(x) g(x) \tag{10}$$

It is easy to check that this satisfies the linearity condition and also induces positive norm (because of the positivity of $w(x)$). For different values of a and b and choices of the 'weight function' w, the Gram-Schmidt orthonormalization leads to different families of orthogonal polynomials (Hermite, Laguerre, Chebyshev, etc.). From the point of view of polynomial approximation, a given weight function could include the information about certain regions where we wish the polynomial approximation to be more accurate. The choice of $w(x) = 1$ treats the entire domain democratically. However, say we wish to be more precise near the middle of the interval. Then, a weight function which is maximum in the middle of the interval will yield the desired approximation.
Of particular importance is the choice $w(x) = 1/\sqrt{1-x^2}$ corresponding to the interval $[-1,1]$. This choice generates the Chebyshev Polynomials, which have the special 'min-max' property, making zeros of these polynomials the preferred choice of interpolation points in the problem of polynomial interpolation.