III.3 Orthogonal Polynomials

Fourier series proved very powerful for approximating periodic functions. If periodicity is lost, however, uniform convergence is lost. In this chapter we introduce alternative bases, *Orthogonal Polynomials (OPs)* built on polynomials that are applicable in the non-periodic setting. That is we consider expansions of the form

$$f(x) = \sum_{k=0}^{\infty} c_k p_k(x)$$

where $p_k(x)$ are special families of polynomials, c_k are expansion coefficients. The computation of the coefficients $c_k \approx c_k^n$ will wait until the next section.

Why not use monomials as in Taylor series? Hidden in the previous lecture was that we could effectively compute Taylor coefficients by evaluating on the unit circle in the complex plane, *only* if the radius of convergence was 1. Many functions are smooth on say $[-1, 1]$ but have non-convergent Taylor series, e.g.:

$$\frac{1}{25x^2 + 1}$$

While orthogonal polynomials span the same space as monomials, and therefore we can in theory write an approximation in monomials, orthogonal polynomials are *much* more stable.

In addition to numerics, OPs play a very important role in many mathematical areas including functional analysis, integrable systems, singular integral equations, complex analysis, and random matrix theory.

1. General properties: we define orthogonal polynomials
2. Three-term recurrences
3. Jacobi matrices

Remark (advanced) Going beyond what we discuss here are many other beautiful properties of orthogonal polynomials that are useful in computation, analysis, representation theory, quantum mechanics, etc. These include sparse recurrence relationships for derivatives, that they are eigenfunctions of differential equations, their asymptotics, generating functions, etc.

1. General properties

Definition 1 (graded polynomial basis) A set of polynomials $\{p_0(x), p_1(x), \ldots\}$ is *graded* if p_n is precisely degree n: i.e.,

$$p_n(x) = k_n x^n + k_n^{(1)} x^{n-1} + \cdots + k_n^{(n-1)} x + k_n^{(n)}$$

for $k_n \neq 0$.

Note that if p_n are graded then $\{p_0(x), \ldots, p_n(x)\}$ are a basis of all polynomials of degree n.

Definition 2 (orthogonal polynomials) Given an (integrable) *weight $w(x) > 0$* for $x \in (a, b)$, which defines a continuous inner product

$$\langle f, g \rangle = \int_a^b f(x) g(x) w(x) \mathrm{d}x$$

a graded polynomial basis $\{p_0(x), p_1(x), \ldots\}$ are *orthogonal polynomials (OPs)* if

$$\langle p_n, p_m \rangle = 0$$

whenever $n \neq m$. We assume through that integrals of polynomials are finite:

$$\int_a^b x^k w(x) \mathrm{d}x < \infty.$$

Note in the above

$$h_n := \langle p_n, p_n \rangle = \|p_n\|^2 = \int_a^b p_n(x)^2 w(x) \mathrm{d}x > 0.$$

Multiplying any orthogonal polynomial by a nonzero constant necessarily is also an orthogonal polynomial. We have two standard normalisations:

Definition 3 (orthonormal polynomials) A set of orthogonal polynomials $\{q_0(x), q_1(x), \ldots\}$ are *orthonormal* if $\|q_n\| = 1$.

Definition 4 (monic orthogonal polynomials) A set of orthogonal polynomials $\{\pi_0(x), \pi_1(x), \ldots\}$ are *monic* if $k_n = 1$.

Proposition 1 (existence) Given a weight $w(x)$, monic orthogonal polynomials exist.

Proof

Existence follows immediately from the Gram–Schmidt procedure. That is, define $\pi_0(x) := 1$ and

$$\pi_n(x) := x^n - \sum_{k=0}^{n-1} \frac{\langle x^n, \pi_k \rangle}{\|\pi_k\|^2} \pi_k(x)$$

∎

We are primarily concerned with the usage of orthogonal polynomials in approximating functions. First we observe the following:

Proposition 2 (expansion) If $r(x)$ is a degree n polynomial and $\{p_n\}$ are orthogonal then

$$r(x) = \sum\nolimits_{k=0}^{n} \frac{\langle p_k, r \rangle}{\|p_k\|^2} p_k(x)$$

Note for $\{q_n\}$ orthonormal we have

$$r(x) = \sum\nolimits_{k=0}^{n} \langle q_k, r \rangle q_k(x).$$

Proof Because $\{p_0, \ldots, p_n\}$ are a basis of polynomials we can write

$$r(x) = \sum\nolimits_{k=0}^{n} r_k p_k(x)$$

for constants $r_k \in \mathbb{R}$. By linearity we have

$$\langle p_m, r \rangle = \sum\nolimits_{k=0}^{n} r_k \langle p_m, p_k \rangle = r_m \langle p_m, p_m \rangle$$

∎

Corollary 1 (zero inner product) If a degree n polynomial r satisfies

$$0 = \langle p_0, r \rangle = \ldots = \langle p_n, r \rangle$$

then $r = 0$.

Corollary 2 (uniqueness) Monic orthogonal polynomials are unique.

Proof If $p_n(x)$ and $\pi_n(x)$ are both monic orthogonal polynomials then $r(x) = p_n(x) - \pi_n(x)$ is degree $n - 1$ but satisfies

$$\langle r, \pi_k \rangle = \langle p_n, \pi_k \rangle - \langle \pi_n, \pi_k \rangle = 0$$

for $k = 0, \ldots, n - 1$. Note $\langle p_n, \pi_k \rangle = 0$ can be seen by expanding

$$\pi_k(x) = \sum\nolimits_{j=0}^{k} c_j p_j(x).$$

∎

OPs are uniquely defined (up to a constant) by the property that they are orthogonal to all lower degree polynomials.

Theorem 1 (orthogonal to lower degree) Given a weight $w(x)$, a polynomial

$$p(x) = k_n x^n + O(x^{n-1})$$

with $k_n \neq 0$ satisfies

$$\langle p, f_m \rangle = 0$$

for all polynomials f_m of degree $m < n$ if and only if $p(x) = k_n \pi_n(x)$ where $\pi_n(x)$ are the monic orthogonal polynomials. Therefore an orthogonal polynomial is uniquely defined by the weight and leading order coefficient k_n.

We leave this proof to the problem sheets. A consequence of this is that orthonormal polynomials are always a constant multiple of orthogonal polynomials.

2. 3-term recurrence

The most *fundamental* property of orthogonal polynomials is their three-term recurrence.

Theorem 2 (3-term recurrence, 2nd form) If $\{p_n\}$ are OPs then there exist real constants $a_n, b_n \neq 0, c_{n-1} \neq 0$ such that

$$xp_0(x) = a_0 p_0(x) + b_0 p_1(x)$$
$$xp_n(x) = c_{n-1}p_{n-1}(x) + a_n p_n(x) + b_n p_{n+1}(x)$$

Proof The $n = 0$ case is immediate since $\{p_0, p_1\}$ are a basis of degree 1 polynomials. The $n > 0$ case follows from

$$\langle xp_n, p_k \rangle = \langle p_n, xp_k \rangle = 0$$

for $k < n - 1$ as xp_k is of degree $k + 1 < n$.

Note that

$$b_n = \frac{\langle p_{n+1}, xp_n \rangle}{\|p_{n+1}\|^2} \neq 0$$

since $xp_n = k_n x^{n+1} + O(x^n)$ is precisely degree n. Further,

$$c_{n-1} = \frac{\langle p_{n-1}, xp_n \rangle}{\|p_{n-1}\|^2} = \frac{\langle p_n, xp_{n-1} \rangle}{\|p_{n-1}\|^2} = b_{n-1} \frac{\|p_n\|^2}{\|p_{n-1}\|^2} \neq 0.$$

∎

Clearly if π_n is monic then so is $x\pi_n$ which leads to the following:

Corollary 3 (monic 3-term recurrence) If $\{\pi_n\}$ are monic then $b_n = 1$.

Corollary 4 (Gram–Schmidt revisited) We can define $\pi_{n+1}(x)$ in terms of π_{n-1} and π_n :

$$\pi_{n+1}(x) = x\pi_n(x) - a_n \pi_n(x) - c_{n-1}\pi_{n-1}(x)$$

where

$$a_n = \frac{\langle x\pi_n, \pi_n \rangle}{\|\pi_n\|^2} \qquad \text{and} \qquad c_{n-1} = \frac{\langle x\pi_n, \pi_{n-1} \rangle}{\|\pi_{n-1}\|^2}$$

Example 1 What are the monic OPs $\pi_0(x), \ldots, \pi_3(x)$ with respect to $w(x) = 1$ on $[0, 1]$? We can construct these using Gram–Schmidt, but exploiting the 3-term recurrence to reduce the computational cost. We have $\pi_0(x) = 1$, which we see is orthogonal:

$$\|\pi_0\|^2 = \langle \pi_0, \pi_0 \rangle = \int_0^1 \mathrm{d}x = 1.$$

We know from the 3-term recurrence that

$$x\pi_0(x) = a_0\pi_0(x) + \pi_1(x)$$

where

$$a_0 = \frac{\langle \pi_0, x\pi_0 \rangle}{\|\pi_0\|^2} = \int_0^1 x\mathrm{d}x = 1/2.$$

Thus

$$\pi_1(x) = x\pi_0(x) - a_0\pi_0(x) = x - 1/2$$
$$\|\pi_1\|^2 = \int_0^1 (x^2 - x + 1/4)\mathrm{d}x = 1/12$$

From

$$x\pi_1(x) = c_0\pi_0(x) + a_1\pi_1(x) + \pi_2(x)$$

we have

$$c_0 = \frac{\langle \pi_0, x\pi_1 \rangle}{\|\pi_0\|^2} = \int_0^1 (x^2 - x/2)\mathrm{d}x = 1/12$$
$$a_1 = \frac{\langle \pi_1, x\pi_1 \rangle}{\|\pi_1\|^2} = 12 \int_0^1 (x^3 - x^2 + x/4)\mathrm{d}x = 1/2$$
$$\pi_2(x) = x\pi_1(x) - c_0 - a_1\pi_1(x) = x^2 - x + 1/6$$
$$\|\pi_2\|^2 = \int_0^1 (x^4 - 2x^3 + 4x^2/3 - x/3 + 1/36)\mathrm{d}x = \frac{1}{180}$$

Finally, from

$$x\pi_2(x) = c_1\pi_1(x) + a_2\pi_2(x) + \pi_3(x)$$

we have

$$c_1 = \frac{\langle \pi_1, x\pi_2 \rangle}{\|\pi_1\|^2} = 12 \int_0^1 (x^4 - 3x^3/2 + 2x^2/3 - x/12)\mathrm{d}x = 1/15$$

$$a_2 = \frac{\langle \pi_2, x\pi_2 \rangle}{\|\pi_2\|^2} = 180 \int_0^1 (x^5 - 2x^4 + 4x^3/3 - x^2/3 + x/36)\mathrm{d}x = 1/2$$

$$\pi_3(x) = x\pi_2(x) - c_1\pi_1(x) - a_2\pi_2(x) = x^3 - x^2 + x/6 - x/15 + 1/30 - x^2/2 + x/2$$
$$= x^3 - 3x^2/2 + 3x/5 - 1/20$$

3. Jacobi matrices

The three-term recurrence can also be interpreted as a matrix known as the Jacobi matrix:

Corollary 5 (Jacobi matrix) For

$$P(x) := [p_0(x)|p_1(x)| \cdots]$$

then we have

$$xP(x) = P(x) \underbrace{\begin{bmatrix} a_0 & c_0 & & & \\ b_0 & a_1 & c_1 & & \\ & b_1 & a_2 & \ddots & \\ & & \ddots & \ddots & \end{bmatrix}}_{X}$$

More generally, for any polynomial $a(x)$ we have

$$a(x)P(x) = P(x)a(X).$$

The Jacobi matrix is $J := X^\top$.

For the special cases of orthonormal and monic polynomials we have extra structure:

Corollary 6 (orthonormal 3-term recurrence) The Jacobi matrix of a family of orthogonal polynomials $p_n(x)$ is symmetric:

$$J = X = \begin{bmatrix} a_0 & b_0 & & & \\ b_0 & a_1 & b_1 & & \\ & b_1 & a_2 & \ddots & \\ & & \ddots & \ddots & \end{bmatrix}$$

if and only if $p_n(x)$ is up-to-sign a fixed constant scaling of orthonormal: for $q_n(x) := \pi_n(x)/\|\pi_n\|^2$ we have for a fixed $\alpha \in \mathbb{R}$ and $s_n \in \{-1, 1\}$

$$p_n(x) = \alpha s_n q_n(x).$$

Proof Noting that $\|q_n\|^2 = 1$ and thence $\|p_n\|^2 = \alpha^2$, if $p_n(x) = \alpha s_n q_n(x)$ we have

$$b_n = \frac{\langle xp_n, p_{n+1}\rangle}{\|p_{n+1}\|^2} = s_n s_{n+1}\langle xq_n, q_{n+1}\rangle = s_n s_{n+1}\langle q_n, xq_{n+1}\rangle = \frac{\langle p_n, xp_{n+1}\rangle}{\|p_n\|^2} = c_{n-1}.$$

Conversely, suppose $X = X^\top$, i.e., $b_n = c_{n-1}$ and write the corresponding orthogonal polynomials as $p_n(x) = \alpha_n q_n(x)$. We have

$$b_n = \frac{\langle xp_n, p_{n+1}\rangle}{\|p_{n+1}\|^2} = \frac{\alpha_n}{\alpha_{n+1}}\langle xq_n, q_{n+1}\rangle = \frac{\alpha_n}{\alpha_{n+1}}\langle q_n, xq_{n+1}\rangle = \frac{\alpha_n^2}{\alpha_{n+1}^2}\frac{\langle p_n, xp_{n+1}\rangle}{\|p_n\|^2} = \frac{\alpha_n^2}{\alpha_{n-}^2}$$

$$= \frac{\alpha_n^2}{\alpha_{n+1}^2}b_n.$$

Hence $\alpha_n^2 = \alpha_{n+1}^2$ which implies that $\alpha_{n+1} = \pm\alpha_n$. By induction the result follows, where $\alpha := \alpha_0$. ∎

Remark (advanced) If you are worried about multiplication of infinite matrices/vectors note it is well-defined by the standard definition because it is banded. It can also be defined in terms of functional analysis where one considers these as linear operators (functions of functions) between vector spaces.

Remark (advanced) Every integrable weight generates a family of orthonormal polynomials, which in turn generates a symmetric Jacobi matrix. There is a "Spectral Theorem for Jacobi matrices" that says one can go the other way: every tridiagonal symmetric matrix with bounded entries is a Jacobi matrix for some integrable weight with compact support. This is an example of what Barry Simon calls a "Gem of spectral theory", that is.

Example 2 (uniform weight Jacobi matrix) Consider the monic orthogonal polynomials $\pi_0(x), \pi_1(x), \ldots, \pi_3(x)$ for $w(x) = 1$ on $[0, 1]$ constructed above. We can write the 3-term recurrence coefficients we have computed above as the Jacobi matrix:

$$x[\pi_0(x)|\pi_1(x)|\cdots] = [\pi_0(x)|\pi_1(x)|\cdots]\underbrace{\begin{bmatrix} 1/2 & 1/12 & & & \\ 1 & 1/2 & 1/15 & & \\ & 1 & 1/2 & \ddots & \\ & & \ddots & \ddots & \end{bmatrix}}_{X}$$

We can compute the orthonormal polynomials, using

$$\|\pi_3\|^2 = \int_0^1 (x^6 - 3x^5 + 69x^4/20 - 19x^3/10 + 51x^2/100 - 3x/50 + 1/400)\mathrm{d}x = \frac{1}{28\mathrm{(}}$$

as:

$$q_0(x) = \pi_0(x)$$
$$q_1(x) = \sqrt{12}\pi_1(x) = \sqrt{3}(2x - 1)$$
$$q_2(x) = \sqrt{180}\pi_2(x) = \sqrt{5}(6x^2 - 6x + 1)$$
$$q_3(x) = \sqrt{2800}\pi_3(x) = \sqrt{7}(20x^3 - 30x^2 + 12x - 1)$$

which have the Jacobi matrix

$$x[q_0(x)|q_1(x)|\cdots] = x[\pi_0(x)|\pi_1(x)|\cdots]\underbrace{\begin{bmatrix} 1 & & & & \\ & 2\sqrt{3} & & & \\ & & 6\sqrt{5} & & \\ & & & 20\sqrt{7} & \\ & & & & \ddots \end{bmatrix}}_{D}$$

$$= [q_0(x)|q_1(x)|\cdots]D^{-1}XD = \begin{bmatrix} 1/2 & 1/\sqrt{12} & & \\ 1/\sqrt{12} & 1/2 & 1/\sqrt{15} & \\ & 1/\sqrt{15} & 1/2 & \ddots \\ & & \ddots & \ddots \end{bmatrix}$$

which is indeed symmetric. The problem sheet explores a more elegant way of doing this.

Example 3 (expansion) Consider expanding a low degree polynomial like $f(x) = x^2$ in $\pi_n(x)$. We have

$$\langle \pi_0, f \rangle = \int_0^1 x^2 \mathrm{d}x = 1/3 \langle \pi_1, f \rangle = \int_0^1 x^2(x - 1/2)\mathrm{d}x = 1/12 \langle \pi_2, f \rangle =$$

$$\int_0^1 x^2(x^2 - x + 1/6)\mathrm{d}x = 1/180$$

Thus we have:

$$f(x) = \frac{\pi_0(x)}{3} + \pi_1(x) + \pi_2(x) = [\pi_0(x)|\pi_1(x)|\pi_2(x)|\cdots]\begin{bmatrix} 1/3 \\ 1 \\ 1 \\ 0 \\ \vdots \end{bmatrix}$$

We multiply (using that $b_2 = 1$ for monic OPs) to deduce:

$$xf(x) = x[\pi_0(x)|\pi_1(x)|\pi_2(x)|\cdots]\begin{bmatrix} 1/3 \\ 1 \\ 1 \\ 0 \\ \vdots \end{bmatrix}$$

$$= [\pi_0(x)|\pi_1(x)|\pi_2(x)|\cdots]X\begin{bmatrix} 1/3 \\ 1 \\ 1 \\ 0 \\ \vdots \end{bmatrix} = [\pi_0(x)|\pi_1(x)|\pi_2(x)|\cdots]\begin{bmatrix} 1/4 \\ 9/10 \\ 3/2 \\ 1 \\ 0 \\ \vdots \end{bmatrix}$$

$$= \frac{\pi_0(x)}{4} + \frac{9\pi_1(x)}{10} + \frac{3\pi_2(x)}{2} + \pi_3(x)$$