Chapter VI

Orthogonal Polynomials

Fourier series are very powerful for approximating periodic functions. If periodicity is lost, however, the Fourier coefficients are no longer in ℓ^1 and uniform convergence is lost. In this chapter we introduce *Orthogonal Polynomials (OPs)* which are more effective for computing with non-periodic (but still continuous/differentiable) functions. 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 and c_k are expansion coefficients. The approximation of the coefficients $c_k \approx c_k^n$ using quadrature will be explored later.

Why not use monomials as in Taylor series? Hidden in the discussion on Fourier series 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, orthogonal polynomials are much more stable and can effectively and accurately approximate such functions. In particular, where we saw that interpolation by monomials at evenly spaced points performed horribly in practice we can use orthogonal polynomials with specially chosen points to get reliable interpolation of functions.

1. VI.1 General Orthogonal Polynomials: For non-periodic functions we consider the definition of orthogonal polynomials, and discuss their basic properties.

2. VI.2 Classical Orthogonal Polynomials: For certain weights, orthogonal polynomials are classical and have addition structure that are useful for computations.

3. VI.3 Gaussian Quadrature: Finally, we revisit the problem of computing integrals, and see that using orthogonal polynomials we can derive much more accurate methods.

We stop at integration, but Fourier and orthogonal polynomial expansions also lead to very effective scheme for solving differential equations and many other applications. 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.

VI.1 General Orthogonal Polynomials

A family of orthogonal polynomials is a special case of a *graded polynomial basis*:

Definition 35 (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 36 (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 37 (Orthonormal Polynomials)**.** A set of orthogonal polynomials $\{q_0(x), q_1(x), \ldots\}$ are *orthonormal* if $\|q_n\| = 1$.

Definition 38 (Monic Orthogonal Polynomials)**.** A set of orthogonal polynomials $\{\pi_0(x), \pi_1(x), \ldots\}$ are *monic* if $k_n = 1$.

Proposition 13 (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

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

Assume π_m are monic OPs for all $m < n$. Then we have

$$\langle \pi_m, \pi_n \rangle = \langle \pi_m, x^n \rangle - \sum_{k=0}^{n-1} \frac{\langle x^n, \pi_k \rangle}{\|\pi_k\|^2} \underbrace{\langle \pi_m, \pi_k \rangle}_{=0 \text{ if } m \neq k} = \langle \pi_m, x^n \rangle - \langle x^n, \pi_m \rangle = 0.$$

■

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

Proposition 14 (expansion)**.** *If $r(x)$ is a degree n polynomial and $\{p_n\}$ are orthogonal then*

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

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

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

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

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

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

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

which implies that $r_m = \langle p_m, r\rangle / \langle p_m, p_m\rangle$. ■

Corollary 5 (zero inner product)**.** *If a degree n polynomial r satisfies*

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

then $r = 0$.

Proof If all the inner products are zero the coefficients in the expansion are all zero and r is zero. ■

Corollary 6 (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

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

■

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

Theorem 14 (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.

Proof We leave this proof to the problem sheets. ■

VI.1.1 Three-term recurrence

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

Theorem 15 (3-term recurrence, 2nd form). *If $\{p_n\}$ are OPs then there exist real constants $a_n b_n, c_{n-1}$ such that*

$$
\begin{aligned}
xp_0(x) &= a_0 p_0(x) + b_0 p_1(x) \\
xp_n(x) &= c_{n-1}p_{n-1}(x) + a_n p_n(x) + b_n p_{n+1}(x),
\end{aligned}
$$

where $b_n \neq 0$ and $c_{n-1} \neq 0$.

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 7 (monic 3-term recurrence). *$\{\pi_n\}$ are monic if and only if $b_n = 1$.*

Proof

If $b_n = 1$ and $\pi_n(x) = x^n + O(x^{n-1})$ then the 3-term recurrence shows us that

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

and $\pi_{n+1}(x)$ is also monic. Similarly, if $\pi_n(x)$ is monic and $b_n \neq 1$ then $\pi_{n+1}(x)$ is not monic, which would be a contradiction. ■

Note this implies that 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 23 (constructing OPs). 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 already normalised:

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

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

$$\begin{aligned} \pi_1(x) &= x\pi_0(x) - a_0\pi_0(x) = x - 1/2 \qquad \Rightarrow \\ \|\pi_1\|^2 &= \int_0^1 (x^2 - x + 1/4)\mathrm{d}x = 1/12. \end{aligned}$$

From

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

we have

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

Finally, from

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

we have

$$\begin{aligned} c_1 &= \frac{\langle \pi_1, x\pi_2 \rangle}{\|\pi_1\|^2} = 12\int_0^1 (x^4 - 3x^3/2 + 2x^2/3 - x/12)\mathrm{d}x = 1/15, \\ a_2 &= \frac{\langle \pi_2, x\pi_2 \rangle}{\|\pi_2\|^2} = 180\int_0^1 (x^5 - 2x^4 + 4x^3/3 - x^2/3 + x/36)\mathrm{d}x = 1/2, \\ \pi_3(x) &= x\pi_2(x) - c_1\pi_1(x) - a_2\pi_2(x) \\ &= x^3 - x^2 + x/6 - x/15 + 1/30 - x^2/2 + x/2 - 1/12 \\ &= x^3 - 3x^2/2 + 3x/5 - 1/20 \end{aligned}$$

VI.1.2 Jacobi matrices

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

Corollary 8 (multiplication 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

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

Proof The expression follows:

$$xP(x) = [xp_0(x)|xp_1(x)|\cdots] = [a_0p_0(x) + b_0p_1(x)|c_0p_0(x) + a_1p_1(x) + b_1p_2(x)|\cdots] = P(x)X.$$

For polynomials, note that

$$x^k P(x) = x^{k-1}P(x)X = \cdots = P(x)X^k.$$

Thus if $a(x) = \sum_{k=0}^n a_k x^k$ we have by linearity

$$a(x)P(x) = \sum_{k=0}^n a_k x^k P(x) = P(x)\sum_{k=0}^n a_k X^k = P(x)a(X).$$

■

Remark 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.

For the special cases of orthonormal polynomials we have extra structure, in which case we refer to the matrix as a *Jacobi matrix*:

Corollary 9 (Jacobi matrix)**.** *The multiplication matrix of a family of orthogonal polynomials $p_n(x)$ is symmetric,*

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

if and only if $p_n(x)$ is up-to-sign a fixed constant scaling of orthonormal: for $q_n(x) := \pi_n(x)/\|\pi_n\|$ 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 First, assume $p_n(x)$ has the specified form. 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

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

and therefore X is symmetric.

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

$$\begin{aligned} b_n &= \frac{\langle xp_n, p_{n+1}\rangle}{\|p_{n+1}\|^2} = \frac{\alpha_n}{\alpha_{n+1}}\langle xq_n, q_{n+1}\rangle = \frac{\alpha_n}{\alpha_{n+1}}\langle q_n, xq_{n+1}\rangle = \frac{\alpha_n^2}{\alpha_{n+1}^2}\frac{\langle p_n, xp_{n+1}\rangle}{\|p_n\|^2} \\ &= \frac{\alpha_n^2}{\alpha_{n+1}^2}c_n = \frac{\alpha_n^2}{\alpha_{n+1}^2}b_n. \end{aligned}$$

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 Every compactly supported integrable weight generates a family of orthonormal polynomials, which in turn generates a 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".

Example 24 (uniform weight orthonormal polynomials)**.** Consider computing orthonormal polynomials with respect to $w(x) = 1$ on $[0, 1]$. Above we constructed the monic OPs $\pi_0(x), \ldots, \pi_3(x)$ so we can deduce the orthonormal polynomials by dividing by their norm, but there is another way: writing $q_n(x) = k_n\pi_n(x)$, find the normalisation k_n that turns the 3-term recurrence into a symmetric matrix. We can write the 3-term recurrence coefficients for monic OPs as a multiplication matrix:

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

The previous theorem says that if we rescale the polynomials so that the resulting Jacobi matrix is symmetric than they are by necessity the orthonormal polynomials. In particular, consider writing:

$$[q_0(x)|q_1(x)|\cdots] = [\pi_0(x)|\pi_1(x)|\cdots]\underbrace{\begin{bmatrix} k_0 & & & \\ & k_1 & & \\ & & k_2 & \\ & & & \ddots \end{bmatrix}}_{K}$$

where we want to find the normalisation constants. Since $\|\pi_0\| = 1$ we know $k_0 = 1$. We have

$$x[q_0(x)|q_1(x)|\cdots] = [\pi_0(x)|\pi_1(x)|\cdots]XK = [q_0(x)|q_1(x)|\cdots]\underbrace{K^{-1}XK}_{J}$$

where

$$J = \begin{bmatrix} a_0 & c_0k_1 & & & \\ k_1^{-1} & a_1 & c_1k_2/k_1 & & \\ & k_1/k_2 & a_2 & c_2k_3/k_2 & \\ & & \ddots & \ddots & \ddots \end{bmatrix}.$$

Thus for this to be symmetric we need

$$c_0k_1 = k_1^{-1}, c_1k_2/k_1 = k_2^{-1}, c_2k_3/k_2 = k_3^{-1}, \ldots$$

Note that

$$c_2 = \frac{\langle \pi_2, x\pi_3\rangle}{\|\pi_2\|^2} = 180\int_0^1 (x^6 - 5x^5/2 + 34x^4/15 - 9x^3/10 + 3x^2/20 - x/120)\mathrm{d}x = 9/140.$$

Thus we have (noting that the k_n are all positive which simplifies the square-roots):

$$\begin{aligned} k_1 &= 1/\sqrt{c_0} = 2\sqrt{3}, \\ k_2 &= k_1/\sqrt{c_1} = 6\sqrt{5}, \\ k_3 &= k_2/\sqrt{c_2} = 20\sqrt{7}. \end{aligned}$$

Thus we have

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

which have the Jacobi matrix

$$J = \begin{bmatrix} 1/2 & 1/(2\sqrt{3}) & & & \\ 1/(2\sqrt{3}) & 1/2 & 1/\sqrt{15} & & \\ & 1/\sqrt{15} & 1/2 & 3/(2\sqrt{35}) & \\ & & 3/(2\sqrt{35}) & 1/2 & \ddots \\ & & & \ddots & \ddots \end{bmatrix}.$$

Example 25 (expansion via Jacobi matrix)**.** What are the expansion coefficients of $x^3 - x + 1$ in $\{q_n\}$? We could deduce this by computing the inner products though its actually simpler to use the multiplication matrix. In particular if we write

$$Q(x) := [q_0(x)|q_1(x)|q_2(x)|\cdots]$$

Then we have (note: $q_0(x) \equiv 1$ only because the weight integrates to 1) $1 = Q(x)\boldsymbol{e}_1$. This tells us that:

$$x = xQ(x)\boldsymbol{e}_1 = Q(x)X\boldsymbol{e}_1 = Q(x)\begin{bmatrix} 1/2 \\ 1/(2\sqrt{3}) \\ 0 \\ \vdots \end{bmatrix}.$$

Continuing we have

$$\begin{aligned} x^2 &= Q(x)X\begin{bmatrix} 1/2 \\ 1/(2\sqrt{3}) \\ 0 \\ \vdots \end{bmatrix} = Q(x)\begin{bmatrix} 1/3 \\ 1/(2\sqrt{3}) \\ 1/(6\sqrt{5}) \\ 0 \\ \vdots \end{bmatrix} \\ x^3 &= Q(x)X\begin{bmatrix} 1/3 \\ 1/(2\sqrt{3}) \\ 1/(6\sqrt{5}) \\ 0 \\ \vdots \end{bmatrix} = Q(x)\begin{bmatrix} 1/4 \\ \frac{3\sqrt{3}}{20} \\ \frac{1}{4\sqrt{5}} \\ \frac{1}{20\sqrt{7}} \\ 0 \\ \vdots \end{bmatrix} \end{aligned}$$

Thus by linearity we find that

$$\begin{aligned} x^3 - x + 1 &= Q(x)\begin{bmatrix} 3/4 \\ -1/(20\sqrt{3}) \\ \frac{1}{4\sqrt{5}} \\ \frac{1}{20\sqrt{7}} \\ 0 \\ \vdots \end{bmatrix} \\ &= \frac{3}{4}q_0(x) - \frac{1}{20\sqrt{3}}q_1(x) + \frac{1}{4\sqrt{5}}q_2(x) + \frac{1}{20\sqrt{7}}q_3(x). \end{aligned}$$

VI.2 Classical Orthogonal Polynomials

Classical orthogonal polynomials are special families of orthogonal polynomials with a number of beautiful properties, for example (1) their derivatives are also OPs and (2) they are eigenfunctions of simple differential operators. As stated above orthogonal polynomials are uniquely defined by the weight $w(x)$ and the constant k_n and hence we can define the classical OPs by specifying their weights and normalisation constants.

The classical orthogonal polynomials are:

1. Chebyshev polynomials (1st kind) $T_n(x)$: $w(x) = 1/\sqrt{1-x^2}$ on $[-1,1]$.
2. Chebyshev polynomials (2nd kind) $U_n(x)$: $\sqrt{1-x^2}$ on $[-1,1]$.
3. Legendre polynomials $P_n(x)$: $w(x) = 1$ on $[-1,1]$.
4. Ultrapsherical polynomials (my fav!): $C_n^{(\lambda)}(x)$: $w(x) = (1-x^2)^{\lambda-1/2}$ on $[-1,1]$, $\lambda \neq 0$, $\lambda > -1/2$.
5. Jacobi polynomials: $P_n^{(a,b)}(x)$: $w(x) = (1-x)^a(1+x)^b$ on $[-1,1]$, $a, b > -1$.
6. Laguerre polynomials: $L_n(x)$: $w(x) = \exp(-x)$ on $[0,\infty)$.
7. Hermite polynomials $H_n(x)$: $w(x) = \exp(-x^2)$ on $(-\infty,\infty)$.

In the notes we will discuss:

1. Chebyshev polynomials: These are closely linked to Fourier series and are one of the most powerful tools in numerics.
2. Legendre polynomials: These have no simple closed-form expression but can be defined in terms of a Rodriguez formula, a feature that applies to all other classical families.

VI.2.1 Chebyshev polynomials

There are four families of Chebyshev polynomials but we will consider the first two:

Definition 39 (Chebyshev polynomials, 1st kind)**.** $T_n(x)$ are orthogonal with respect to $1/\sqrt{1-x^2}$ and satisfy:

$$\begin{aligned} T_0(x) &= 1, \\ T_n(x) &= 2^{n-1}x^n + O(x^{n-1}) \end{aligned}$$

Definition 40 (Chebyshev polynomials, 2nd kind)**.** $U_n(x)$ are orthogonal with respect to $\sqrt{1-x^2}$ and satisfy:

$$U_n(x) = 2^n x^n + O(x^{n-1})$$

A beautiful fact is that Chebyshev polynomials are really trigonometric polynomials in disguise:

Theorem 16 (Chebyshev T are cos)**.** *For* $-1 \leq x \leq 1$

$$T_n(x) = \cos n \operatorname{acos} x.$$

In other words

$$T_n(\cos\theta) = \cos n\theta.$$

Proof

We need to show that $p_n(x) := \cos n\mathrm{acos}\, x$ are

1. graded polynomials
2. orthogonal w.r.t. $1/\sqrt{1-x^2}$ on $[-1, 1]$, and
3. have the right normalisation constant $k_n = 2^{n-1}$ for $n = 2, \ldots$.

Property (2) follows under a change of variables:

$$\int_{-1}^{1} \frac{p_n(x)p_m(x)}{\sqrt{1-x^2}} \mathrm{d}x = \int_0^\pi \frac{\cos(n\theta)\cos(m\theta)}{\sqrt{1-\cos^2\theta}} \sin\theta \mathrm{d}\theta = \int_0^\pi \cos(n\theta)\cos(m\theta)\mathrm{d}x = 0$$

if $n \neq m$.

To see that they are graded we use the fact that

$$xp_n(x) = \cos\theta\cos n\theta = \frac{\cos(n-1)\theta + \cos(n+1)\theta}{2} = \frac{p_{n-1}(x) + p_{n+1}(x)}{2}.$$

In other words $p_{n+1}(x) = 2xp_n(x) - p_{n-1}(x)$. Since each time we multiply by $2x$ and $p_0(x) = 1$ we have by induction

$$p_{n+1}(x) = 2x(2^{n-1}x^n + O(x^{n-1})) + O(x^{n-1})) = 2^n x^{n+1} + O(x^n)$$

which completes the proof.

■

Recall that the 3-term recurrence is an important property of a family of orthogonal polynomials. We can deduce from the relationship with cosines the following:

Corollary 10 (Chebyshev 3-term recurrence)**.**

$$\begin{aligned} xT_0(x) &= T_1(x) \\ xT_n(x) &= \frac{T_{n-1}(x) + T_{n+1}(x)}{2} \end{aligned}$$

Proof This is rewriting the expression we used to show that $p_n(x)$ are graded in the previous proof. ■

Chebyshev polynomials are particularly powerful as their expansions are cosine series in disguise: for

$$f(x) = \sum_{k=0}^{\infty} \check{f}_k T_k(x)$$

we have

$$f(\cos\theta) = \sum_{k=0}^{\infty} \check{f}_k \cos k\theta.$$

Thus the coefficients can be recovered fast using FFT-based techniques.

We will also see the following:

Theorem 17 (Chebyshev U are sin)**.** *For $x = \cos\theta$,*

$$U_n(x) = \frac{\sin(n+1)\theta}{\sin\theta}$$

which satisfy:

$$\begin{aligned} xU_0(x) &= U_1(x)/2 \\ xU_n(x) &= \frac{U_{n-1}(x)}{2} + \frac{U_{n+1}(x)}{2}. \end{aligned}$$

Proof Shown in the problem sheet. ■

VI.2.2 Legendre polynomials

Definition 41 (Legendre)**.** Legendre polynomials $P_n(x)$ are orthogonal polynomials with respect to $w(x) = 1$ on $[-1, 1]$, with

$$k_n = \frac{1}{2^n}\binom{2n}{n} = \frac{(2n)!}{2^n (n!)^2}$$

The reason for this complicated normalisation constant is both historical and that it leads to simpler formulae for recurrence relationships.

Classical orthogonal polynomials have *Rodriguez formulae*, defining orthogonal polynomials as high order derivatives of simple functions. In this case we have:

Lemma 9 (Legendre Rodriguez formula)**.**

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

Proof We need to verify:

1. graded polynomials
2. orthogonal to all lower degree polynomials on $[-1, 1]$, and
3. have the right normalisation constant $k_n = \frac{1}{2^n}\binom{2n}{n}$.

(1) follows since its a degree n polynomial (the n-th derivative of a degree $2n$ polynomial). (2) follows by integration by parts. Note that $(1-x^2)^n$ and its first $n-1$ derivatives vanish at ± 1. If r_m is a degree $m < n$ polynomial we have:

$$\begin{aligned}\int_{-1}^{1} \frac{\mathrm{d}^n}{\mathrm{d}x^n}(1-x^2)^n r_m(x)\mathrm{d}x &= -\int_{-1}^{1} \frac{\mathrm{d}^{n-1}}{\mathrm{d}x^{n-1}}(1-x^2)^n r_m'(x)\mathrm{d}x \\ &= \cdots = (-)^n \int_{-1}^{1}(1-x^2)^n r_m^{(n)}(x)\mathrm{d}x = 0.\end{aligned}$$

(3) follows since:

$$\begin{aligned}\frac{\mathrm{d}^n}{\mathrm{d}x^n}[(-)^n x^{2n} + O(x^{2n-1})] &= (-)^n 2n \frac{\mathrm{d}^{n-1}}{\mathrm{d}x^{n-1}} x^{2n-1} + O(x^{2n-1})] \\ &= (-)^n 2n(2n-1)\frac{\mathrm{d}^{n-2}}{\mathrm{d}x^{n-2}} x^{2n-2} + O(x^{2n-2})] = \cdots \\ &= (-)^n 2n(2n-1)\cdots(n+1)x^n + O(x^{n-1}) \\ &= (-)^n \frac{(2n)!}{n!} x^n + O(x^{n-1})\end{aligned}$$

■

This allows us to determine the coefficients $k_n^{(\lambda)}$ which are useful in proofs. In particular we will use $k_n^{(2)}$:

Lemma 10 (Legendre monomial coefficients)**.**

$$\begin{aligned}P_0(x) &= 1 \\ P_1(x) &= x \\ P_n(x) &= \underbrace{\frac{(2n)!}{2^n(n!)^2}}_{k_n} x^n - \underbrace{\frac{(2n-2)!}{2^n(n-2)!(n-1)!}}_{k_n^{(2)}} x^{n-2} + O(x^{n-4}).\end{aligned}$$

Here the $O(x^{n-4})$ is as $x \to \infty$, which implies that the term is a polynomial of degree $\leq n-4$. For $n = 2, 3$ the $O(x^{n-4})$ term is therefore precisely zero.

Proof

The $n = 0$ and 1 case are immediate. For the other case we expand $(1-x^2)^n$ to get:

$$\begin{aligned}(-)^n \frac{\mathrm{d}^n}{\mathrm{d}x^n}(1-x^2)^n &= \frac{\mathrm{d}^n}{\mathrm{d}x^n}[x^{2n} - nx^{2n-2} + O(x^{2n-4})] \\ &= (2n)\cdots(2n-n+1)x^n - n(2n-2)\cdots(2n-2-n+1)x^{n-2} + O(x^{n-4}) \\ &= \frac{(2n)!}{n!}x^n - \frac{n(2n-2)!}{(n-2)!}x^{n-2} + O(x^{n-4})\end{aligned}$$

Multiplying through by $\frac{1}{2^n(n!)}$ completes the derivation.

■

Theorem 18 (Legendre 3-term recurrence)**.**

$$\begin{aligned} xP_0(x) &= P_1(x) \\ (2n+1)xP_n(x) &= nP_{n-1}(x) + (n+1)P_{n+1}(x) \end{aligned}$$

Proof The $n = 0$ case is immediate (since $w(x) = w(-x)$ we know $a_n = 0$, from the problem sheet). For the other cases we match terms:

$$\begin{aligned} (2n+1)&xP_n(x) - nP_{n-1}(x) - (n+1)P_{n+1}(x) \\ &= [(2n+1)k_n - (n+1)k_{n+1}]x^{n+1} \\ &\quad + [(2n+1)k_n^{(2)} - nk_{n-1} - (n+1)k_{n+1}^{(2)}]x^{n-1} + O(x^{n-3}) \end{aligned}$$

Using the expressions for k_n and $k_n^{(2)}$ above we have (leaving the manipulations as an exercise):

$$\begin{aligned} (2n+1)k_n - (n+1)k_{n+1} &= \frac{(2n+1)!}{2^n(n!)^2} - (n+1)\frac{(2n+2)!}{2^{n+1}((n+1)!)^2} = 0 \\ (2n+1)k_n^{(2)} - nk_{n-1} - (n+1)k_{n+1}^{(2)} &= -(2n+1)\frac{(2n-2)!}{2^n(n-2)!(n-1)!} - n\frac{(2n-2)!}{2^{n-1}((n-1)!)^2} \\ &\quad + (n+1)\frac{(2n)!}{2^{n+1}(n-1)!n!} \\ &= 0 \end{aligned}$$

Thus

$$(2n+1)xP_n(x) - nP_{n-1}(x) - (n+1)P_{n+1}(x) = O(x^{n-3})$$

But as it is orthogonal to $P_k(x)$ for $0 \leq k \leq n-3$ it must be zero. ■

VI.3 Gaussian Quadrature

We have already seen examples of quadrature including the Rectangular, Trapezium and Interpolatory Quadrature Rules. In this section we see that a special type of interpolatory quadrature can be constructed by using the roots of orthogonal polynomials, leading to a method that is exact for polynomials of twice the expected degree. Importantly, we can use quadrature to compute expansions in orthogonal polynomials that interpolate, mirroring the link between the Trapezium rule, Fourier series, and interpolation but now for orthogonal polynomials.

We begin with a simple example demonstrating the power of using roots of orthogonal polynomials in an interpolatory quadrature rule:

Example 26 (Gauss-Chebyshev)**.** We find the interpolatory quadrature rule for $w(x) = 1/\sqrt{1-x^2}$ on $[-1, 1]$ with points equal to the roots of $T_3(x)$. This is a special case of Gaussian quadrature which we will approach in another way below. We use:

$$\int_{-1}^1 w(x)\mathrm{d}x = \pi, \int_{-1}^1 xw(x)\mathrm{d}x = 0, \int_{-1}^1 x^2w(x)\mathrm{d}x = \frac{\pi}{2}.$$

From the 3-term recurrence we deduce

$$T_0(x) = 1, T_1(x) = x, T_2(x) = 2xT_1(x) - T_0(x) = 2x^2 - 1, T_3(x) = 2xT_2(x) - T_1(x) = 4x^3 - 3x$$

hence we find the 3 roots of $T_3(x)$ are $x_1, x_2, x_3 = \sqrt{3}/2, 0, -\sqrt{3}/2$. Thus we have:

$$\begin{aligned}
w_1 &= \int_{-1}^{1} w(x)\ell_1(x)\mathrm{d}x = \int_{-1}^{1} \frac{x(x+\sqrt{3}/2)}{(\sqrt{3}/2)\sqrt{3}\sqrt{1-x^2}}\mathrm{d}x = \frac{\pi}{3} \\
w_2 &= \int_{-1}^{1} w(x)\ell_2(x)\mathrm{d}x = \int_{-1}^{1} \frac{(x-\sqrt{3}/2)(x+\sqrt{3}/2)}{(-3/4)\sqrt{1-x^2}}\mathrm{d}x = \frac{\pi}{3} \\
w_3 &= \int_{-1}^{1} w(x)\ell_3(x)\mathrm{d}x = \int_{-1}^{1} \frac{(x-\sqrt{3}/2)x}{(-\sqrt{3})(-\sqrt{3}/2)\sqrt{1-x^2}}\mathrm{d}x = \frac{\pi}{3}
\end{aligned}$$

(It's not a coincidence that they are all the same but this will differ for roots of other OPs.) That is we have

$$\Sigma_n^{w,\boldsymbol{x}}[f] = \frac{\pi}{3}\left[f(\sqrt{3}/2) + f(0) + f(-\sqrt{3}/2)\right].$$

This is indeed exact for polynomials up to degree $n - 1 = 2$, but it goes all the way up to $2n - 1 = 5$:

$$\begin{aligned}
&\Sigma_n^{w,\boldsymbol{x}}[1] = \pi, \Sigma_n^{w,\boldsymbol{x}}[x] = 0, \Sigma_n^{w,\boldsymbol{x}}[x^2] = \frac{\pi}{2}, \\
&\Sigma_n^{w,\boldsymbol{x}}[x^3] = 0, \Sigma_n^{w,\boldsymbol{x}}[x^4] = \frac{3\pi}{8}, \Sigma_n^{w,\boldsymbol{x}}[x^5] = 0 \\
&\Sigma_n^{w,\boldsymbol{x}}[x^6] = \frac{9\pi}{32} \neq \frac{5\pi}{16}
\end{aligned}$$

We shall explain this miracle in the rest of this section.

VI.3.1 Roots of orthogonal polynomials and truncated Jacobi matrices

Consider roots (zeros) of orthogonal polynomials $p_n(x)$ which will be essential to constructing Gaussian quadrature, via interpolation at these points. For interpolation to be well-defined we first need to guarantee that the roots are distinct.

Lemma 11 (OP roots)**.** *An orthogonal polynomial $p_n(x)$ has exactly n distinct roots.*

Proof

Suppose $x_1, \ldots, x_j$ are the roots where $p_n(x)$ changes sign, i.e., the order of the root must be odd and hence

$$p_n(x) = c_k(x - x_k)^{2p+1} + O((x - x_k)^{2p+2})$$

for $c_k \neq 0$ and $k = 1, \ldots, j$ and $p \in \mathbb{Z}$, as $x \to x_k$. Then

$$p_n(x)(x - x_1)\cdots(x - x_j)$$

does not change signs: it behaves like $c_k(x - x_k)^{2p+2} + O(x - x_k)^{2p+3}$ as $x \to x_k$. In other words:

$$\langle p_n, (x - x_1)\cdots(x - x_j)\rangle = \int_a^b p_n(x)(x - x_1)\cdots(x - x_j)w(x)\mathrm{d}x \neq 0.$$

where $w(x)$ is the weight of orthogonality. This is only possible if $j = n$ as $p_n(x)$ is orthogonal w.r.t. all lower degree polynomials and hence otherwise this integral would be zero. Since $p_n(x)$ is exactly degree n it follows each root must be first order and hence distinct.

■

We can relate these roots to eigenvalues of truncations of Jacobi matrices:

Definition 42 (truncated Jacobi matrix)**.** Given a Jacobi matrix J associated with a family of orthonormal polynomials, the *truncated Jacobi matrix* is

$$J_n := \begin{bmatrix} a_0 & b_0 & & \\ b_0 & \ddots & \ddots & \\ & \ddots & a_{n-2} & b_{n-2} \\ & & b_{n-2} & a_{n-1} \end{bmatrix} \in \mathbb{R}^{n\times n}$$

Lemma 12 (OP roots and Jacobi matrices)**.** *The zeros $x_1, \ldots, x_n$ of an orthonormal polynomial $q_n(x)$ are the eigenvalues of the truncated Jacobi matrix J_n. More precisely,*

$$J_n Q_n = Q_n \begin{bmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{bmatrix}$$

for the orthogonal matrix

$$Q_n = \underbrace{\begin{bmatrix} q_0(x_1) & \cdots & q_0(x_n) \\ \vdots & \cdots & \vdots \\ q_{n-1}(x_1) & \cdots & q_{n-1}(x_n) \end{bmatrix}}_{V_n^\top} \begin{bmatrix} \alpha_1^{-1} & & \\ & \ddots & \\ & & \alpha_n^{-1} \end{bmatrix}$$

where $\alpha_j = \sqrt{q_0(x_j)^2 + \cdots + q_{n-1}(x_j)^2}$.

Proof

We construct the eigenvector (noting $b_{n-1}q_n(x_j) = 0$):

$$J_n \begin{bmatrix} q_0(x_j) \\ \vdots \\ q_{n-1}(x_j) \end{bmatrix} = \begin{bmatrix} a_0 q_0(x_j) + b_0 q_1(x_j) \\ b_0 q_0(x_j) + a_1 q_1(x_j) + b_1 q_2(x_j) \\ \vdots \\ b_{n-3} q_{n-3}(x_j) + a_{n-2} q_{n-2}(x_j) + b_{n-2} q_{n-1}(x_j) \\ b_{n-2} q_{n-2}(x_j) + a_{n-1} q_{n-1}(x_j) + b_{n-1} q_n(x_j) \end{bmatrix} = x_j \begin{bmatrix} q_0(x_j) \\ q_1(x_j) \\ \vdots \\ q_{n-1}(x_j) \end{bmatrix}$$

The spectral theorem guarantees that all symmetric matrices have an orthogonal eigenvector matrix. That is, by scaling the columns of the eigenvectors we know there must exist α_j so that

$$Q_n = \underbrace{\begin{bmatrix} q_0(x_1) & \cdots & q_0(x_n) \\ \vdots & \cdots & \vdots \\ q_{n-1}(x_1) & \cdots & q_{n-1}(x_n) \end{bmatrix}}_{V_n^\top} \begin{bmatrix} \alpha_1^{-1} & & \\ & \ddots & \\ & & \alpha_n^{-1} \end{bmatrix}$$

is orthogonal. We choose α_j so that

$$\boldsymbol{e}_j^\top Q_n^\top Q_n \boldsymbol{e}_j = \frac{\sum_{k=0}^{n-1} q_k(x_j)^2}{\alpha_j^2} = 1.$$

■

Example 27 (Chebyshev roots)**.** Consider $T_n(x) = \cos n\operatorname{acos} x$. The roots are $x_j = \cos\theta_j$ where $\theta_j = (j - 1/2)\pi/n$ for $j = 1, \ldots, n$ are the roots of $\cos n\theta$ that are inside $[0, \pi]$.

Consider the $n = 3$ case where we have

$$x_1, x_2, x_3 = \cos(\pi/6), \cos(\pi/2), \cos(5\pi/6) = \sqrt{3}/2, 0, -\sqrt{3}/2$$

We also have from the 3-term recurrence:

$$\begin{aligned}
T_0(x) &= 1, \\
T_1(x) &= x, \\
T_2(x) &= 2xT_1(x) - T_0(x) = 2x^2 - 1, \\
T_3(x) &= 2xT_2(x) - T_1(x) = 4x^3 - 3x.
\end{aligned}$$

As determined as part of the problem sheet, we orthonormalise by rescaling

$$\begin{aligned}
q_0(x) &= 1/\sqrt{\pi}, \\
q_k(x) &= T_k(x)\sqrt{2}/\sqrt{\pi}.
\end{aligned}$$

so that the Jacobi matrix is symmetric:

$$x[q_0(x)|q_1(x)|\cdots] = [q_0(x)|q_1(x)|\cdots] \underbrace{\begin{bmatrix} 0 & 1/\sqrt{2} & & & \\ 1/\sqrt{2} & 0 & 1/2 & & \\ & 1/2 & 0 & 1/2 & \\ & & 1/2 & 0 & \ddots \\ & & & \ddots & \ddots \end{bmatrix}}_{J}$$

We can then confirm that we have constructed an eigenvector/eigenvalue of the 3×3 truncation of the Jacobi matrix, e.g. at $x_2 = 0$:

$$\begin{bmatrix} 0 & 1/\sqrt{2} & \\ 1/\sqrt{2} & 0 & 1/2 \\ & 1/2 & 0 \end{bmatrix} \begin{bmatrix} q_0(0) \\ q_1(0) \\ q_2(0) \end{bmatrix} = \frac{1}{\sqrt{\pi}} \begin{bmatrix} 0 & 1/\sqrt{2} & \\ 1/\sqrt{2} & 0 & 1/2 \\ & 1/2 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -\sqrt{2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

VI.3.2 Properties of Gaussian quadrature

We now introduce Gaussian quadrature, which we shall see is exact for polynomials up to degree $2n - 1$, i.e., double the degree of other interpolatory quadrature rules from other grids.

Rather than defining Gaussian quadrature as an interpolatory quadrature rule, we build an analogy with the discrete Fourier transform (DFT) by defining a quadrature rule for which our orthogonal polynomials satisfy a discrete orthogonality property.

Definition 43 (Gaussian quadrature)**.** Given a weight $w(x)$, the Gauss quadrature rule is:

$$\int_a^b f(x)w(x)\mathrm{d}x \approx \underbrace{\sum_{j=1}^n w_j f(x_j)}_{\Sigma_n^w[f]}$$

where $x_1, \ldots, x_n$ are the roots of the orthonormal polynomials $q_n(x)$ and

$$w_j := \frac{1}{\alpha_j^2} = \frac{1}{q_0(x_j)^2 + \cdots + q_{n-1}(x_j)^2}.$$

Equivalentally, $x_1, \ldots, x_n$ are the eigenvalues of J_n and w_j can be written in terms of the eigenvectors and the integral of the weight:

$$w_j = \int_a^b w(x)\mathrm{d}x \underbrace{q_0(x_j)^2/\alpha_j^2}_{Q_n[1,j]^2}.$$

In analogy to how Fourier series are orthogonal with respect to the Trapezium rule, Orthogonal polynomials are orthogonal with respect to Gaussian quadrature:

Lemma 13 (Discrete orthogonality)**.** *For $0 \le \ell, m \le n-1$, the orthonormal polynomials $q_n(x)$ satisfy*

$$\Sigma_n^w[q_\ell q_m] = \delta_{\ell m}$$

Proof

$$\Sigma_n^w[q_\ell q_m] = \sum_{j=1}^n \frac{q_\ell(x_j)q_m(x_j)}{\alpha_j^2} = \left[\frac{q_\ell(x_1)}{\alpha_1}|\cdots|\frac{q_\ell(x_n)}{\alpha_n}\right]\begin{bmatrix} q_m(x_1)/\alpha_1 \\ \vdots \\ q_m(x_n)/\alpha_n \end{bmatrix} = \boldsymbol{e}_\ell^\top Q_n Q_n^\top \boldsymbol{e}_m = \delta_{\ell m}$$

■

Just as approximating Fourier coefficients using Trapezium rule gives a way of interpolating at the grid, so does Gaussian quadrature:

Theorem 19 (interpolation via quadrature)**.** *For the orthonormal polynomials $q_n(x)$,*

$$f_n(x) := \sum_{k=0}^{n-1} c_k^n q_k(x) \text{ for } c_k^n := \Sigma_n^w[f q_k]$$

interpolates $f(x)$ at the Gaussian quadrature points $x_1, \ldots, x_n$.

Proof Consider the Vandermonde-like matrix from above:

$$V_n := \begin{bmatrix} q_0(x_1) & \cdots & q_{n-1}(x_1) \\ \vdots & \ddots & \vdots \\ q_0(x_n) & \cdots & q_{n-1}(x_n) \end{bmatrix}$$

and define

$$Q_n^w := V_n^\top \begin{bmatrix} w_1 & & \\ & \ddots & \\ & & w_n \end{bmatrix} = \begin{bmatrix} q_0(x_1)w_1 & \cdots & q_0(x_n)w_n \\ \vdots & \ddots & \vdots \\ q_{n-1}(x_1)w_1 & \cdots & q_{n-1}(x_n)w_n \end{bmatrix}$$

so that

$$\begin{bmatrix} c_0^n \\ \vdots \\ c_{n-1}^n \end{bmatrix} = Q_n^w \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{bmatrix}.$$

Note that if $p(x) = [q_0(x)|\cdots|q_{n-1}(x)]\boldsymbol{c}$ then

$$\begin{bmatrix} p(x_1) \\ \vdots \\ p(x_n) \end{bmatrix} = V_n \boldsymbol{c}$$

But we see that (similar to the Fourier case)

$$Q_n^w V_n = \begin{bmatrix} \Sigma_n^w[q_0 q_0] & \cdots & \Sigma_n^w[q_0 q_{n-1}] \\ \vdots & \ddots & \vdots \\ \Sigma_n^w[q_{n-1} q_0] & \cdots & \Sigma_n^w[q_{n-1} q_{n-1}] \end{bmatrix} = I$$

and hence $V_n^{-1} = Q_n^w$ and we have

$$f_n(x_j) = [q_0(x_j)|\cdots|q_{n-1}(x_j)] Q_n^w \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{bmatrix} = \boldsymbol{e}_j^\top V_n Q_n^w \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{bmatrix} = f(x_j).$$

■

Example 28 (Chebyshev expansions)**.** Consider the construction of Gaussian quadrature associated with the Chebyshev weight for $n = 3$. To determine the weights we need we compute

$$w_j^{-1} = \alpha_j^2 = q_0(x_j)^2 + q_1(x_j)^2 + q_2(x_j)^2 = \frac{1}{\pi} + \frac{2}{\pi} x_j^2 + \frac{2}{\pi}(2x_j^2 - 1)^2$$

We can check each case and deduce that $w_j = \pi/3$. Thus we recover the interpolatory quadrature rule. Further, we can construct the transform

$$\begin{aligned} Q_3^w &= \begin{bmatrix} w_1 q_0(x_1) & w_2 q_0(x_2) & w_3 q_0(x_3) \\ w_1 q_1(x_1) & w_2 q_1(x_2) & w_3 q_1(x_3) \\ w_1 q_2(x_1) & w_2 q_2(x_2) & w_3 q_2(x_3) \end{bmatrix} \\ &= \frac{\pi}{3} \begin{bmatrix} 1/\sqrt{\pi} & 1/\sqrt{\pi} & 1/\sqrt{\pi} \\ x_1\sqrt{2/\pi} & x_2\sqrt{2/\pi} & x_3\sqrt{2/\pi} \\ (2x_1^2 - 1)\sqrt{2/\pi} & (2x_2^2 - 1)\sqrt{2/\pi} & (2x_3^2 - 1)\sqrt{2/\pi} \end{bmatrix} \\ &= \frac{\sqrt{\pi}}{3} \begin{bmatrix} 1 & 1 & 1 \\ \sqrt{6}/2 & 0 & -\sqrt{6}/2 \\ 1/\sqrt{2} & -\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \end{aligned}$$

We can use this to expand a polynomial, e.g. x^2:

$$Q_3^w \begin{bmatrix} x_1^2 \\ x_2^2 \\ x_3^2 \end{bmatrix} = \frac{\sqrt{\pi}}{3} \begin{bmatrix} 1 & 1 & 1 \\ \sqrt{6}/2 & 0 & -\sqrt{6}/2 \\ 1/\sqrt{2} & -\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 3/4 \\ 0 \\ 3/4 \end{bmatrix} = \begin{bmatrix} \sqrt{\pi}/2 \\ 0 \\ \sqrt{\pi}/(2\sqrt{2}) \end{bmatrix}$$

In other words:

$$x^2 = \frac{\sqrt{\pi}}{2} q_0(x) + \frac{\sqrt{\pi}}{2\sqrt{2}} q_2(x) = \frac{1}{2} T_0(x) + \frac{1}{2} T_2(x)$$

which can be easily confirmed.

Corollary 11 (Gaussian quadrature is interpolatory)**.** *Gaussian quadrature is an interpolatory quadrature rule with the interpolation points equal to the roots of q_n:*

$$\Sigma_n^w[f] = \int_a^b f_n(x)w(x)\mathrm{d}x.$$

Proof We want to show that its the same as integrating the interpolatory polynomial:

$$\int_a^b f_n(x)w(x)\mathrm{d}x = \frac{1}{q_0(x)} \sum_{k=0}^{n-1} c_k^n \int_a^b q_k(x)q_0(x)w(x)\mathrm{d}x = \frac{c_0^n}{q_0} = \Sigma_n^w[f].$$

■

A consequence of being an interpolatory quadrature rule is that it is exact for all polynomials of degree $n-1$. The *miracle* of Gaussian quadrature is it is exact for twice as many!

Theorem 20 (Exactness of Gauss quadrature)**.** *If $p(x)$ is a degree $2n-1$ polynomial then Gauss quadrature is exact:*

$$\int_a^b p(x)w(x)\mathrm{d}x = \Sigma_n^w[p].$$

Proof Using polynomial division algorithm (e.g. by matching terms) we can write

$$p(x) = q_n(x)s(x) + r(x)$$

where s and r are degree $n-1$ and $q_n(x)$ is the degree n orthonormal polynomial. Because Gauss quadrature is interpolatory we know that it is exact for degree $n-1$ polynomials, in particular:

$$\Sigma_n^w[r] = \int_a^b r(x)w(x)\mathrm{d}x.$$

But then we find that

$$\begin{aligned}
\Sigma_n^w[p] &= \underbrace{\Sigma_n^w[q_n s]}_{\text{0 since evaluating } q_n \text{ at zeros}} + \Sigma_n^w[r] = \int_a^b r(x)w(x)\mathrm{d}x \\
&= \underbrace{\int_a^b q_n(x)s(x)w(x)\mathrm{d}x}_{\text{0 since } s \text{ is degree } < n} + \int_a^b r(x)w(x)\mathrm{d}x \\
&= \int_a^b p(x)w(x)\mathrm{d}x.
\end{aligned}$$

■

Example 29 (Double exactness)**.** Let's look at an example in completeness for $n = 3$ with uniform weight on $[-1, 1]$. From the 3-term recurrence for Legendre polynomials we get the multiplication matrix

$$x[P_0(x)|P_1(x)|\cdots] = [P_0(x)|P_1(x)|\cdots] \underbrace{\begin{bmatrix} 0 & 1/3 & & & \\ 1 & 0 & 2/5 & & \\ & 2/3 & 0 & 3/7 & \\ & & 3/5 & 0 & \ddots \\ & & & \ddots & \ddots \end{bmatrix}}_{X}$$

From this we deduce that

$$\begin{aligned}
P_0(x) &= 1 \\
P_1(x) &= x \\
P_2(x) &= 3/2xP_1(x) - P_0(x)/2 = \frac{3x^2}{2} - \frac{1}{2} \\
P_3(x) &= 5/3xP_2(x) - 2P_1(x)/3 = \frac{5x^2}{2} - \frac{3x}{2}.
\end{aligned}$$

The roots of $P_3(x)$ are

$$x_1, x_2, x_3 = -\sqrt{3/5}, 0, \sqrt{3/5}.$$

We know the first orthonormal polynomial is $q_0(x) = 1/\sqrt{2}$, i.e., $k_0 = 1/\sqrt{2}$. We write

$$[q_0(x)|q_1(x)|\cdots] = [P_0(x)|P_1(x)|\cdots] \underbrace{\begin{bmatrix} 1/\sqrt{2} & & & \\ & k_1 & & \\ & & k_2 & \\ & & & \ddots \end{bmatrix}}_{K}$$

Thus from

$$x[q_0(x)|q_1(x)|\cdots] = [q_0(x)|q_1(x)|\cdots] \underbrace{K^{-1}XK}_{J}$$

we find that

$$J = \begin{bmatrix} 0 & \sqrt{2}k_1/3 & & & \\ 1/(\sqrt{2}k_1) & 0 & 2k_2/(5k_1) & & \\ & 2k_1/(3k_2) & 0 & 3k_3/(7k_2) & \\ & & 3k_2/(5k_3) & 0 & \ddots \\ & & & \ddots & \ddots \end{bmatrix}$$

For this to be symmetric we find

$$\begin{aligned}
k_1 &= \sqrt{3/2} \\
k_2 &= \sqrt{10k_1^2/6} = \sqrt{5/2} \\
k_3 &= \sqrt{21k_2^2/15} = \sqrt{21/6}
\end{aligned}$$

We thus get the quadrature weights

$$\begin{aligned}
w_1 &= \alpha_1^{-2} = \frac{1}{q_0(x_1)^2 + q_1(x_1)^2 + q_2(x_1)^2} = \frac{1}{1/2 + (3/2)\times(3/5) + (5/2)\times(4/25)} = \frac{5}{9} \\
w_2 &= \alpha_2^{-2} = \frac{1}{q_0(x_2)^2 + q_1(x_2)^2 + q_2(x_2)^2} = \frac{1}{1/2 + (5/2)\times(1/4)} = \frac{8}{9} \\
w_3 &= w_1 = \frac{5}{9}.
\end{aligned}$$

Thus our Gauss–Legendre quadrature formula is

$$\Sigma_3^w[f] = \frac{5}{9}f(-\sqrt{3/5}) + \frac{8}{9}f(0) + \frac{5}{9}f(\sqrt{3/5}).$$

We are exact for all polynomials up to degree $2n-1=5$:

$$\begin{aligned}
\Sigma_3^w[1] &= \frac{5}{9}+\frac{8}{9}+\frac{5}{9}=2\\
\Sigma_3^w[x] &= -\frac{5}{9}\sqrt{3/5}+\frac{5}{9}\sqrt{3/5}=0\\
\Sigma_3^w[x^2] &= \frac{5}{9}\frac{3}{5}+\frac{5}{9}\frac{3}{5}=\frac{2}{3}\\
\Sigma_3^w[x^3] &= -\frac{5}{9}(3/5)^{3/2}+\frac{5}{9}(3/5)^{3/2}=0\\
\Sigma_3^w[x^4] &= \frac{5}{9}\frac{9}{25}+\frac{5}{9}\frac{9}{25}=\frac{2}{5}\\
\Sigma_3^w[x^5] &= 0.
\end{aligned}$$

But the next integral is wrong:

$$\Sigma_3^w[x^6]=\frac{5}{9}\frac{27}{125}+\frac{5}{9}\frac{27}{125}=\frac{6}{25}\neq\frac{2}{7}=\int_{-1}^{1}x^6\mathrm{d}x.$$

Going beyond polynomials, Gaussian quadrature achieves faster than algebraic convergence for any smooth function. If the function is analytic in a neighbourhood of the support of the interval this is in fact exponential convergence, far exceeding the convergence rate observed for rectangular and Trapezium rules. This is a beautiful example of more sophisticated mathematics leading to powerful numerical methods.