

Mathematical Tripos Part IB: Lent Term 2024

Numerical Analysis – Lecture 3

3 Orthogonal polynomials

3.1 Three-term recurrence relation

Consider $\mathbb{X} = C[a, b]$, the space of all continuous real-valued functions $f : [a, b] \rightarrow \mathbb{R}$, and define a scalar (or inner) product on $C[a, b]$ by

$$(f, g) := (f, g)_w := \int_a^b f(x)g(x)w(x)dx. \quad (3.1)$$

Here w , the so-called *weight function*, is a fixed nonnegative function, such that the integral $\int w(x)dx$ exists (then $\int h(x)w(x)dx$ exists for all $h \in C[a, b]$).

We denote by $\mathcal{P}_n$ the space of all algebraic polynomials of degree (at most) n , i.e., $p \in \mathcal{P}_n$ if $p(x) = \sum_{k=0}^n a_k x^k$. If the *leading* coefficient a_n equals 1, then p is called a *monic* polynomial. Given a scalar product (3.1), we say that $Q_n \in \mathcal{P}_n$ is the n -th *orthogonal polynomial* if

$$(Q_n, p) = 0 \quad \forall p \in \mathcal{P}_{n-1}.$$

Note: different weights lead to different orthogonal polynomials.

Lemma 3.1 *For every $n \in \mathbb{N}$, there exists a unique monic orthogonal polynomial $Q_n \in \mathcal{P}_n$. Any $p \in \mathcal{P}_n$ is uniquely expressible as a linear combination*

$$p = \sum_{k=0}^n c_k Q_k, \quad c_k = \frac{(p, Q_k)}{(Q_k, Q_k)}. \quad (3.2)$$

Proof. We apply the so-called *Gram-Schmidt orthogonalization* algorithm to the linearly independent sequence of monomials $(1, x, \dots, x^n, \dots)$. Starting with $Q_0 \equiv 1$ we set

$$Q_n(x) := x^n - \sum_{k=0}^{n-1} \frac{(x^n, Q_k)}{(Q_k, Q_k)} Q_k(x), \quad n = 1, 2, \dots$$

Then, from this construction, $(Q_n, Q_k) = 0$ for all $k = \overline{0, n-1}$, i.e., Q_n is orthogonal to any previous Q_k . Also, we have $x^n \in \text{span}(Q_k)_{k=0}^n$, hence $\mathcal{P}_n = \text{span}(Q_k)_{k=0}^n$. Therefore $Q_n \perp \mathcal{P}_{n-1}$, and any $p \in \mathcal{P}_n$ has an expansion (3.2). Each coefficient c_k in (3.2) is uniquely determined by multiplying both sides (in the scalar product sense) with Q_k .

If $\tilde{Q}_n$ is another n -th monic orthogonal polynomial, then $p := Q_n - \tilde{Q}_n$ belongs to $\mathcal{P}_{n-1}$, therefore $(p, p) = (Q_n - \tilde{Q}_n, p) = 0$. Hence $p \equiv 0$ (by the scalar product property), i.e., $\tilde{Q}_n \equiv Q_n$. $\square$

Remark 3.2 For practical construction of orthogonal polynomials the Gram-Schmidt is not of much help, for it leads to loss of accuracy due to imprecisions in calculation of scalar products. A considerably better procedure is being provided by the next theorem.

Theorem 3.3 (Three-term recurrence relation) *Monic orthogonal polynomials satisfy the relation*

$$Q_{n+1}(x) = (x - a_n)Q_n(x) - b_n Q_{n-1}(x), \quad n = 0, 1, \dots \quad (3.3)$$

where $Q_{-1}(x) \equiv 0$, $Q_0(x) \equiv 1$, and

$$a_n = \frac{(xQ_n, Q_n)}{(Q_n, Q_n)}, \quad b_n = \frac{(Q_n, Q_n)}{(Q_{n-1}, Q_{n-1})} > 0. \quad (3.4)$$

Proof. Based on (3.2), let us look at the coefficients of the expansion of $p(x) = xQ_n(x) \in \mathcal{P}_{n+1}$,

$$xQ_n(x) = \sum_{k=0}^{n+1} c_k Q_k(x), \quad c_k \stackrel{(3.2)}{=} \frac{(xQ_n, Q_k)}{(Q_k, Q_k)} \stackrel{(*)}{=} \frac{(Q_n, xQ_k)}{(Q_k, Q_k)}.$$

To find the values of c_k , we will use the last (equivalent) formula (*).

$k = n + 1 \rightarrow c_{n+1} = 1$, since both $xQ_n(x)$ and $Q_{n+1}(x)$ are monic polynomials of degree $n + 1$.

$k = n \rightarrow c_n = \frac{(Q_n, xQ_n)}{(Q_n, Q_n)} = a_n$ in (3.4).

$k = n - 1 \rightarrow$ Because of monicity, we have the equality $xQ_{n-1} = Q_n + p_{n-1}$ where $p_{n-1} \in \mathcal{P}_{n-1}$,
thus $(Q_n, xQ_{n-1}) = (Q_n, Q_n + p_{n-1}) = (Q_n, Q_n)$, hence $c_{n-1} = \frac{(Q_n, Q_n)}{(Q_{n-1}, Q_{n-1})} = b_n$ in (3.4).

$k < n - 1 \rightarrow$ Then $xQ_k \in \mathcal{P}_{n-1}$, and we obtain $(Q_n, xQ_k) = 0$ in (*), thus $c_k = 0$.

It follows that $xQ_n(x) = Q_{n+1} + a_n Q_n + b_n Q_{n-1}$, and that is equivalent to (3.3)-(3.4) $\square$

Remark 3.4 If (Q_k) have leading coefficients (α_k) , then the recurrence takes the form

$$Q_{n+1}(x) = \frac{\alpha_{n+1}}{\alpha_n}(x - a_n)Q_n(x) - \frac{\alpha_{n+1}\alpha_{n-1}}{\alpha_n^2}b_n Q_{n-1}(x), \quad n = 0, 1, \dots$$

and, with an appropriate choice of (α_k) , may become very simple.

3.2 Examples

Example 3.5 Classical examples of orthogonal (non-monic) polynomials include

Name	Notation	Interval	Weight function	Recurrence
Legendre	P_n	$[-1, 1]$	$w(x) \equiv 1$	$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x)$
Chebyshev	T_n	$[-1, 1]$	$w(x) = (1 - x^2)^{-1/2}$	$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$
Laguerre	L_n	$[0, \infty)$	$w(x) = e^{-x}$	$(n+1)L_{n+1}(x) = (2n+1-x)L_n(x) - nL_{n-1}(x)$
Hermite	H_n	$(-\infty, \infty)$	$w(x) = e^{-x^2}$	$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x)$

Example 3.6 (The Chebyshev polynomials) The Chebyshev polynomials T_n , on line 2 in the table above, were introduced in Lecture 2 as

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

where we also proved their three-term recurrence relation in (2.5). Let us show that they are indeed orthogonal with the weight $w(x) = (1 - x^2)^{-1/2}$. With substitution $x = \cos \theta$, $\theta \in [0, \pi]$, we have $T_n(x) = \cos n\theta$, whence

$$\begin{aligned} (T_n, T_m)_w &:= \int_{-1}^1 T_n(x)T_m(x) \frac{dx}{\sqrt{1-x^2}} = \int_0^\pi \cos n\theta \cos m\theta d\theta = \frac{1}{2} \int_{-\pi}^\pi \cos n\theta \cos m\theta d\theta \\ &= \frac{1}{4} \int_{-\pi}^\pi [\cos(n+m)\theta + \cos(n-m)\theta] d\theta = 0, \quad n \neq m. \end{aligned}$$

3.3 Least squares polynomial fitting

Let f be a continuous function defined on some interval $[a, b]$, and suppose we wish to approximate f by a polynomial of degree n . If we equip $C[a, b]$ with the distance $\|f - g\| := (f - g, f - g)^{1/2}$ induced by a scalar product $(f, g) = \int_a^b f(x)g(x)w(x)dx$, then it is natural to seek a polynomial

$$p^* := \arg \min_{p \in \mathcal{P}_n} \|f - p\|,$$

for which the distance $\|f - p^*\|$ is as small as possible. Such a polynomial is called a (*weighted*) *least squares approximant*.

Theorem 3.7 Let $(Q_k)_{k=0}^n$ be polynomials orthogonal with respect to a given inner product. Then the least squares approximant to any $f \in C[a, b]$ from $\mathcal{P}_n$ is given by the formula

$$p^* = p^*(f) = \sum_{k=0}^n c_k^* Q_k, \quad c_k^* = \frac{(f, Q_k)}{\|Q_k\|^2}, \quad (3.5)$$

and the value of the least squares approximation is

$$\|f - p^*\|^2 = \|f\|^2 - \sum_{k=1}^n \frac{(f, Q_k)^2}{\|Q_k\|^2}. \quad (3.6)$$

Remark 3.8 From (3.5), by orthogonality of (Q_k) , the norm of the extremal p^* is equal to

$$\|p^*\|^2 = \left\| \sum_{k=0}^n c_k^* Q_k \right\|^2 = \sum_{k=0}^n |c_k^*|^2 \|Q_k\|^2 = \sum_{k=1}^n \frac{(f, Q_k)^2}{\|Q_k\|^2},$$

hence formula (3.6) takes the form

$$\|f - p^*\|^2 = \|f\|^2 - \|p^*\|^2 \Leftrightarrow \|f - p^*\|^2 + \|p^*\|^2 = \|f\|^2,$$

which is an analogue of the Pythagoras theorem.

Proof. Let $p = \sum_{k=0}^n c_k Q_k$. Then for the squared distance $\|f - p\|^2$ we may write

$$F(c) := (f - p, f - p) = (f - \sum_{k=0}^n c_k Q_k, f - \sum_{k=0}^n c_k Q_k) = \|f\|^2 - 2 \sum_{k=0}^n c_k (f, Q_k) + \sum_{k=0}^n c_k^2 \|Q_k\|^2. \quad (3.7)$$

The right-hand side is a quadratic polynomial in each c_k , therefore it attains its minimum when

$$\left. \frac{\partial F}{\partial c_k} \right|_{c_k = c_k^*} = -2(f, Q_k) + 2c_k^* \|Q_k\|^2 = 0, \quad k = \overline{0, n},$$

hence conclusion (3.5) for c_k^* . Putting optimal c_k^* into (3.7) we obtain (3.6).

Method 3.9 Suppose we want to approximate $f \in C[a, b]$ with a prescribed accuracy ϵ , i.e., to find $n = n(\epsilon)$ such that

$$\|f - p^*\| \leq \epsilon, \quad p^* \in \mathcal{P}_n.$$

From (3.6), it follows that the required value n can be calculated by summing the terms $\frac{(f, Q_k)^2}{\|Q_k\|^2}$ until we reach the bound

$$\sum_{k=1}^n \frac{(f, Q_k)^2}{\|Q_k\|^2} \geq \|f\|^2 - \epsilon^2.$$

The next theorem assures that this bound can be reached indeed.

Theorem 3.10 (The Parseval identity) If $[a, b]$ is finite, then $\sum_{k=0}^{\infty} \frac{(f, Q_k)^2}{\|Q_k\|^2} = \|f\|^2$.

Proof (using the Weierstrass theorem) Let

$$\sigma_n^2 := \|p^*\|^2 = \sum_{k=0}^n \frac{(f, Q_k)^2}{\|Q_k\|^2},$$

so that from (3.5)-(3.6) we have

$$\inf_{p \in \mathcal{P}_n} \|f - p\|^2 = \|f - p^*\|^2 = \|f\|^2 - \sigma_n^2.$$

According to the *Weierstrass theorem*, any function in $C[a, b]$ can be approximated arbitrarily close by a polynomial in the max-norm, hence in the integral norm by means of the inequality

$$\int_a^b |f - p|^2 w(x) dx \leq \max_{x \in [a, b]} |f(x) - p(x)|^2 \int_a^b w(x) dx < \epsilon \cdot \text{const}.$$

Thus $\lim_{n \rightarrow \infty} \|f - p^*\|^2 = \lim_{n \rightarrow \infty} \inf_{p \in \mathcal{P}_n} \|f - p\|^2 = 0$, and we deduce that $\sigma_n^2 \rightarrow \|f\|^2$ as $n \rightarrow \infty$. $\square$

Remark 3.11 Analogous identity is well-known for the trigonometric system (and it is actually valid for any *complete orthogonal system*).