Chapter 8

Orthogonal Polynomials

Given a positive (except possibly at finitely many points), Riemann integrable *weight function* $w(x)$ on $[\,a, b\,]$, the expression

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

defines an inner product on $C[\,a, b\,]$ and

$$\|f\|_2 \;=\; \left(\int_a^b f(x)^2\, w(x)\, dx\right)^{1/2} \;=\; \sqrt{\langle f, f\rangle}$$

defines a *strictly convex* norm on $C[\,a, b\,]$. (See Problem 1 at the end of the chapter.) Thus, given a finite dimensional subspace E of $C[\,a, b\,]$ and an element $f \in C[\,a, b\,]$, there is a unique $g \in E$ such that

$$\|f - g\|_2 \;=\; \min_{h \in E} \|f - h\|_2.$$

We say that g is the *least-squares approximation* to f out of E (relative to w).

Now if we apply the Gram-Schmidt procedure to the sequence $1, x, x^2, \ldots$, we will arrive at a sequence (Q_n) of *orthogonal polynomials* relative to the above inner product. In this special case, however, the Gram-Schmidt procedure simplifies substantially:

Theorem 8.1. *The following procedure defines a sequence (Q_n) of orthogonal polynomials (relative to w). Set:*

$$Q_0(x) = 1, \quad Q_1(x) = x - a_0 = (x - a_0)Q_0(x),$$

and

$$Q_{n+1}(x) = (x - a_n)Q_n(x) - b_n Q_{n-1}(x),$$

for $n \ge 1$, where

$$a_n = \langle\, x\, Q_n, Q_n\,\rangle \big/ \langle\, Q_n, Q_n\,\rangle \quad \textit{and} \quad b_n = \langle\, x\, Q_n, Q_{n-1}\,\rangle \big/ \langle\, Q_{n-1}, Q_{n-1}\,\rangle$$

(and where $x\, Q_n$ is shorthand for the polynomial $x\, Q_n(x)$).

Proof. It's easy to see from these formulas that Q_n is a monic polynomial of degree exactly n. In particular, the Q_n are linearly independent (and nonzero).

Now it's easy to see that Q_0, Q_1, and Q_2 are mutually orthogonal, so let's use induction and check that Q_{n+1} is orthogonal to each Q_k, $k \le n$. First,

$$\langle Q_{n+1}, Q_n \rangle = \langle x\, Q_n, Q_n \rangle - a_n \langle Q_n, Q_n \rangle - b_n \langle Q_{n-1}, Q_n \rangle = 0$$

and

$$\langle Q_{n+1}, Q_{n-1} \rangle = \langle x\, Q_n, Q_{n-1} \rangle - a_n \langle Q_n, Q_{n-1} \rangle - b_n \langle Q_{n-1}, Q_{n-1} \rangle = 0,$$

because $\langle Q_{n-1}, Q_n \rangle = 0$. Next, we take $k < n-1$ and use the recurrence formula twice:

$$\begin{aligned} \langle Q_{n+1}, Q_k \rangle &= \langle x\, Q_n, Q_k \rangle - a_n \langle Q_n, Q_k \rangle - b_n \langle Q_{n-1}, Q_k \rangle \\ &= \langle x\, Q_n, Q_k \rangle = \langle Q_n, x\, Q_k \rangle \qquad \text{(Why?)} \\ &= \langle Q_n, Q_{k+1} + a_k Q_k + b_k Q_{k-1} \rangle = 0, \end{aligned}$$

because $k+1 < n$. □

Remarks 8.2.

1. Using the same trick as above, we have

$$b_n = \langle x\, Q_n, Q_{n-1} \rangle \big/ \langle Q_{n-1}, Q_{n-1} \rangle = \langle Q_n, Q_n \rangle \big/ \langle Q_{n-1}, Q_{n-1} \rangle > 0.$$

2. Each $p \in \mathcal{P}_n$ can be uniquely written $p = \sum_{i=0}^{n} \alpha_i Q_i$, where $\alpha_i = \langle p, Q_i \rangle \big/ \langle Q_i, Q_i \rangle$.

3. If Q is any monic polynomial of degree exactly n, then $Q = Q_n + \sum_{i=0}^{n-1} \alpha_i Q_i$ (why?) and hence

$$\|Q\|_2^2 \;=\; \|Q_n\|_2^2 \;+\; \sum_{i=0}^{n-1} \alpha_i^2 \|Q_i\|_2^2 \;>\; \|Q_n\|_2^2,$$

unless $Q = Q_n$. That is, Q_n has the least $\|\cdot\|_2$ norm of all monic polynomials of degree n.

4. The Q_n are *unique* in the following sense: If (P_n) is another sequence of orthogonal polynomials such that P_n has degree exactly n, then $P_n = \alpha_n Q_n$ for some $\alpha_n \neq 0$. (See Problem 4.) Consequently, there's no harm in referring to the Q_n as *the* sequence of orthogonal polynomials relative to w.

5. For $n \ge 1$ note that $\int_a^b Q_n(t)\, w(t)\, dt = \langle Q_0, Q_n \rangle = 0$.

Examples 8.3.

1. On $[-1, 1]$, the Chebyshev polynomials of the first kind (T_n) are orthogonal relative to the weight $w(x) = 1/\sqrt{1-x^2}$.

$$\begin{aligned} \int_{-1}^{1} T_m(x)\, T_n(x)\, \frac{dx}{\sqrt{1-x^2}} &= \int_0^{\pi} \cos m\theta\, \cos n\theta\, d\theta \\ &= \begin{cases} 0, & m \neq n \\ \pi, & m = n = 0 \\ \pi/2, & m = n \neq 0. \end{cases} \end{aligned}$$

Because T_n has degree exactly n, this must be the right choice. Notice, too, that $\frac{1}{\sqrt{2}}\,T_0, T_1, T_2, \ldots$ are *orthonormal* relative to the weight $2/\pi\sqrt{1-x^2}$.

In terms of the inductive procedure given above, we must have $Q_0 = T_0 = 1$ and $Q_n = 2^{-n+1}T_n$ for $n \ge 1$. (Why?) From this it follows that $a_n = 0$, $b_1 = 1/2$, and $b_n = 1/4$ for $n \ge 2$. (Why?) That is, the recurrence formula given in Theorem 8.1 reduces to the familar relationship $T_{n+1}(x) = 2x\,T_n(x) - T_{n-1}(x)$. Curiously, $Q_n = 2^{-n+1}T_n$ minimizes both

$$\max_{-1\le x\le 1} |p(x)| \qquad \text{and} \qquad \left(\int_{-1}^{1} p(x)^2\,\frac{dx}{\sqrt{1-x^2}}\right)^{1/2}$$

over all monic polynomials of degree exactly n.

The Chebyshev polynomials also satisfy $(1-x^2)\,T_n''(x) - x\,T_n'(x) + n^2\,T_n(x) = 0$. Because this is a polynomial identity, it suffices to check it for all $x = \cos\theta$. In this case,

$$T_n'(x) \;=\; \frac{n\sin n\theta}{\sin\theta}$$

and

$$T_n''(x) \;=\; \frac{n^2\cos n\theta\sin\theta - n\sin n\theta\cos\theta}{\sin^2\theta\,(-\sin\theta)}.$$

Hence,

$$\begin{aligned}&(1-x^2)\,T_n''(x) - x\,T_n'(x) + n^2\,T_n(x)\\ &\qquad = \; -n^2\cos n\theta + n\sin n\theta\cot\theta - n\sin n\theta\cot\theta + n^2\cos\theta = 0\end{aligned}$$

2. On $[-1,1\,]$, the Chebyshev polynomials of the second kind (U_n) are orthogonal relative to the weight $w(x) = \sqrt{1-x^2}$. Indeed,

$$\begin{aligned}&\int_{-1}^{1} U_m(x)\,U_n(x)\,(1-x^2)\,\frac{dx}{\sqrt{1-x^2}}\\ &\qquad = \int_0^{\pi} \frac{\sin\,(m+1)\theta}{\sin\theta}\cdot\frac{\sin\,(n+1)\theta}{\sin\theta}\cdot\sin^2\theta\,d\theta = \begin{cases} 0, & m\ne n\\ \pi/2, & m = n.\end{cases}\end{aligned}$$

While we're at it, notice that

$$T_n'(x) \;=\; \frac{n\sin n\theta}{\sin\theta} \;=\; n\,U_{n-1}(x).$$

As a rule, the derivatives of a sequence of orthogonal polynomials are again orthogonal polynomials, but relative to a different weight.

3. On $[-1,1\,]$ with weight $w(x) \equiv 1$, the sequence (P_n) of *Legendre polynomials* are orthogonal, and are typically normalized by $P_n(1) = 1$. The first few Legendre polynomials are $P_0(x) = 1$, $P_1(x) = x$, $P_2(x) = \frac{3}{2}\,x^2 - \frac{1}{2}$, and $P_3(x) = \frac{5}{2}\,x^3 - \frac{3}{2}\,x$. (Check this!) After we've seen a few more examples, we'll come back and give an explicit formula for P_n.

4. All of the examples we've seen so far are special cases of the following: On $[-1,1]$, consider the weight $w(x) = (1-x)^\alpha(1+x)^\beta$, where $\alpha, \beta > -1$. The corresponding orthogonal polynomials $(P_n^{(\alpha,\beta)})$ are called the *Jacobi polynomials* and are typically normalized by requiring that

$$P_n^{(\alpha,\beta)}(1) \;=\; \binom{n+\alpha}{\alpha} \;\equiv\; \frac{(\alpha+1)(\alpha+2)\cdots(\alpha+n)}{n!}.$$

It follows that $P_n^{(0,0)} = P_n$,

$$P_n^{(-1/2,-1/2)} \;=\; \frac{1\cdot 3\cdot 5\cdots(2n-1)}{2^n n!}\, T_n,$$

and

$$P_n^{(1/2,1/2)} \;=\; \frac{1\cdot 3\cdot 5\cdots(2n+1)}{2^n (n+1)!}\, U_n.$$

The polynomials $P_n^{(\alpha,\alpha)}$ are called *ultraspherical* polynomials.

5. There are also several classical examples of orthogonal polynomials on unbounded intervals. In particular,

$(0,\infty)$	$w(x) = e^{-x}$	Laguerre polynomials,
$(0,\infty)$	$w(x) = x^\alpha e^{-x}$	generalized Laguerre polynomials,
$(-\infty,\infty)$	$w(x) = e^{-x^2}$	Hermite polynomials.

Because Q_n is orthogonal to every element of $\mathcal{P}_{n-1}$, a fuller understanding of Q_n will follow from a characterization of the *orthogonal complement* of $\mathcal{P}_{n-1}$. We begin with an easy fact about least-squares approximations in inner product spaces.

Lemma 8.4. *Let E be a finite dimensional subspace of an inner product space X, and let $x \in X \setminus E$. Then, $y^* \in E$ is the least-squares approximation to x out of E (a.k.a. the nearest point to x in E) if and only if $\langle x - y^*,\, y\rangle = 0$ for every $y \in E$; that is, if and only if $(x - y^*) \perp E$.*

Proof. [We've taken E to be finite dimensional so that nearest points will exist; because X is an inner product space, nearest points must also be unique (see Problem 1 for a proof that every inner product norm is strictly convex).]

($\Longleftarrow$) First suppose that $(x-y^*) \perp E$. Then, given any $y \in E$, we have

$$\|x-y\|_2^2 \;=\; \|(x-y^*) + (y^*-y)\|_2^2 \;=\; \|x-y^*\|_2^2 + \|y^*-y\|_2^2,$$

because $y^* - y \in E$ and, hence, $(x-y^*) \perp (y^*-y)$. Thus, $\|x-y\| > \|x-y^*\|$ unless $y = y^*$; that is, y^* is the (unique) nearest point to x in E.

($\Longrightarrow$) Suppose that $x - y^*$ is *not* orthogonal to E. Then there is some $y \in E$ with $\|y\| = 1$ such that $\alpha = \langle x-y^*,\, y\rangle \neq 0$. It now follows that $y^* + \alpha y \in E$ is a better approximation to x than y^* (and $y^* + \alpha y \neq y^*$, of course); that is, y^* is *not* the least-squares approximation to x. To see this, we again compute:

$$\begin{aligned}\|x-(y^*+\alpha y)\|_2^2 &= \|(x-y^*)-\alpha y\|_2^2 \;=\; \langle (x-y^*)-\alpha y,\, (x-y^*)-\alpha y\rangle \\ &= \|x-y^*\|_2^2 - 2\alpha\,\langle x-y^*,\, y\rangle + \alpha^2 \\ &= \|x-y^*\|_2^2 - \alpha^2 \;<\; \|x-y^*\|_2^2.\end{aligned}$$

Thus, we must have $\langle x-y^*,\, y\rangle = 0$ for every $y \in E$. □

Lemma 8.5. (Integration by-parts)

$$\int_a^b u^{(n)} v \;=\; \sum_{k=1}^{n} (-1)^{k-1} u^{(n-k)} v^{(k-1)} \Big]_a^b \;+\; (-1)^n \int_a^b u v^{(n)}.$$

Now if v is a polynomial of degree $< n$, then $v^{(n)} = 0$ and we get:

Lemma 8.6. *$f \in C[\,a,b\,]$ satisfies $\displaystyle\int_a^b f(x)\,p(x)\,w(x)\,dx = 0$ for all polynomials $p \in \mathcal{P}_{n-1}$ if and only if there is an n-times differentiable function u on $[\,a,b\,]$ satisfying $fw = u^{(n)}$ and $u^{(k)}(a) = u^{(k)}(b) = 0$ for all $k = 0, 1, \ldots, n-1$.*

Proof. One direction is clear from Lemma 8.5: Given u as above, we would have $\int_a^b fpw = \int_a^b u^{(n)} p = (-1)^n \int_a^b u p^{(n)} = 0$.

So, suppose we have that $\int_a^b fpw = 0$ for all $p \in \mathcal{P}_{n-1}$. By integrating fw repeatedly, choosing constants appropriately, we may define a function u satisfying $fw = u^{(n)}$ and $u^{(k)}(a) = 0$ for all $k = 0, 1, \ldots, n-1$. We want to show that the hypotheses on f force $u^{(k)}(b) = 0$ for all $k = 0, 1, \ldots, n-1$.

Now Lemma 8.5 tells us that

$$0 \;=\; \int_a^b fpw \;=\; \sum_{k=1}^{n} (-1)^{k-1} u^{(n-k)}(b)\, p^{(k-1)}(b)$$

for all $p \in \mathcal{P}_{n-1}$. But the numbers $p(b), p'(b), \ldots, p^{(n-1)}(b)$ are completely arbitrary; that is (again by integrating repeatedly, choosing our constants as we please), we can find polynomials p_k of degree $k < n$ such that $p_k^{(k)}(b) \neq 0$ and $p_k^{(j)}(b) = 0$ for $j \neq k$. In fact, $p_k(x) = (x-b)^k$ works just fine! In any case, we must have $u^{(k)}(b) = 0$ for all $k = 0, 1, \ldots, n-1$. □

Rolle's theorem tells us a bit more about the functions orthogonal to $\mathcal{P}_{n-1}$:

Lemma 8.7. *If $w(x) > 0$ in (a, b), and if $f \in C[\,a,b\,]$ is in the orthogonal complement of $\mathcal{P}_{n-1}$ (relative to w); that is, if f satisfies $\int_a^b f(x)\,p(x)\,w(x)\,dx = 0$ for all polynomials $p \in \mathcal{P}_{n-1}$, then f has at least n distinct zeros in the open interval (a, b).*

Proof. Write $fw = u^{(n)}$, where $u^{(k)}(a) = u^{(k)}(b) = 0$ for all $k = 0, 1, \ldots, n-1$. In particular, because $u(a) = u(b) = 0$, Rolle's theorem tells us that u' would have at least one zero in (a, b). But then $u'(a) = u'(c) = u'(b) = 0$, and so u'' must have at least two zeros in (a, b). Continuing, we find that $fw = u^{(n)}$ must have at least n zeros in (a, b). Because $w > 0$, the result follows. □

Corollary 8.8. *Let (Q_n) be the sequence of orthogonal polynomials associated to a given weight w with $w > 0$ in (a, b). Then, the roots of Q_n are real, simple, and lie in (a, b).*

The sheer volume of literature on orthogonal polynomials and other "special functions" is truly staggering. We'll content ourselves with the Legendre and the Chebyshev polynomials. In particular, let's return to the problem of finding an explicit formula for the Legendre polynomials. We could, as Rivlin does, use induction and a few observations that simplify the basic recurrence formula (you're encouraged to read this; see [45, pp. 53–54]). Instead we'll give a simple (but at first sight intimidating) formula that is of use in more general settings than ours.

Lemma 8.6 (with $w \equiv 1$ and $[\,a,b\,] = [-1,1\,]$) says that if we want to find a *polynomial* f of degree n which is orthogonal to $\mathcal{P}_{n-1}$, then we'll need to take a *polynomial* for u, and this u will have to be divisible by $(x-1)^n(x+1)^n$. (Why?) That is, we must have $P_n(x) = c_n \cdot D^n\big[(x^2-1)^n\big]$, where D denotes differentiation, and where c_n is chosen so that $P_n(1) = 1$.

Lemma 8.9. (Leibniz's formula) $D^n(fg) \;=\; \sum_{k=0}^{n} \binom{n}{k} D^k(f)\, D^{n-k}(g).$

Proof. Induction and the fact that $\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k}$. □

Consequently, $Q(x) = D^n\big[(x-1)^n(x+1)^n\big] = \sum_{k=0}^{n}\binom{n}{k} D^k(x-1)^n\, D^{n-k}(x+1)^n$ and it follows that $Q(1) = 2^n n!$ and $Q(-1) = (-1)^n 2^n n!$. This, finally, gives us the formula discovered by Rodrigues in 1814:

$$P_n(x) \;=\; \frac{1}{2^n n!}\, D^n\big[(x^2-1)^n\big]. \tag{8.1}$$

The Rodrigues formula is quite useful (and easily generalizes to the Jacobi polynomials).

Remarks 8.10.

1. By Corollary 8.8, the roots of P_n are real, distinct, and lie in $(-1,1)$.
2. From the binomial theorem, $(x^2-1)^n = \sum_{k=0}^{n}(-1)^k\binom{n}{k}x^{2n-2k}$. If we apply $\frac{1}{2^n n!}\,D^n$ and simplify, we get another formula for the Legendre polynomials.

$$P_n(x) \;=\; \frac{1}{2^n}\sum_{k=0}^{[n/2]}(-1)^k\binom{n}{k}\binom{2n-2k}{n}\,x^{n-2k}.$$

In particular, if n is even (odd), then P_n is even (odd). Notice, too, that if we let $\widetilde{P}_n$ denote the polynomial given by the standard construction, then we must have $P_n = 2^{-n}\binom{2n}{n}\widetilde{P}_n$.

3. In terms of our standard recurrence formula, it follows that $a_n = 0$ (because $xP_n(x)^2$ is always odd). It remains to compute b_n. First, integrating by parts,

$$\int_{-1}^{1} P_n(x)^2\,dx \;=\; xP_n(x)^2\Big]_{-1}^{1} \;-\; \int_{-1}^{1} x\cdot 2P_n(x)\,P_n'(x)\,dx,$$

or $\langle\, P_n, P_n\,\rangle \;=\; 2 - 2\langle\, P_n, xP_n'\,\rangle$. But $xP_n' \;=\; nP_n$ + lower degree terms; hence, $\langle\, P_n, xP_n'\,\rangle \;=\; n\langle\, P_n, P_n\,\rangle$. Thus, $\langle\, P_n, P_n\,\rangle \;=\; 2/(2n+1)$. Using this and the fact that $P_n = 2^{-n}\binom{2n}{n}\widetilde{P}_n$, we'd find that $b_n = n^2/(4n^2-1)$. Thus,

$$\begin{aligned} P_{n+1} \;&=\; 2^{-n-1}\binom{2n+2}{n+1}\widetilde{P}_{n+1} \\ &=\; 2^{-n-1}\binom{2n+2}{n+1}\left[x\,\widetilde{P}_n \;-\; \frac{n^2}{(4n^2-1)}\,\widetilde{P}_{n-1}\right] \\ &=\; \frac{2n+1}{n+1}\,x\,P_n \;-\; \frac{n}{n+1}\,P_{n-1}. \end{aligned}$$

That is, the Legendre polynomials satisfy the recurrence formula

$$(n+1)\,P_{n+1}(x) \;=\; (2n+1)\,x\,P_n(x) - n\,P_{n-1}(x).$$

4. It follows from the calculations in remark 3, above, that the sequence $\widehat{P}_n = \sqrt{\frac{2n+1}{2}}\, P_n$ is *orthonormal* on $[-1, 1\,]$.

5. The Legendre polynomials satisfy the differential equation $(1-x^2)\, P_n''(x) - 2x\, P_n'(x) + n\,(n+1)\, P_n(x) = 0$. If we set $u = (x^2-1)^n$; that is, if $u^{(n)} = 2^n n! P_n$, note that $(x^2-1)u' = 2nxu$. Now we apply D^{n+1} to both sides of this last equation (using Leibniz's formula) and simplify:

$$\begin{aligned} u^{(n+2)}(x^2-1) + (n+1)\, u^{(n+1)}\, 2x + \tfrac{(n+1)n}{2}\, u^{(n)}\, 2 \\ = 2n\left[u^{(n+1)}\, x + (n+1)\, u^{(n)}\right] \\ \implies (1-x^2)\, u^{(n+2)} - 2x\, u^{(n+1)} + n\,(n+1)\, u^{(n)} = 0. \end{aligned}$$

6. Through a series of exercises, similar in spirit to remark 5, Rivlin shows that $|P_n(x)| \le 1$ on $[-1, 1\,]$. See [45, pp. 63–64] for details.

Given an orthogonal sequence, it makes sense to consider *generalized Fourier series* relative to the sequence and to find analogues of the Dirichlet kernel, Lebesgue's theorem, and so on. In case of the Legendre polynomials we have the following:

Example 8.11. The Fourier-Legendre series for $f \in C[-1, 1\,]$ is given by $\sum_k \langle\, f, \widehat{P}_k\,\rangle\, \widehat{P}_k$, where

$$\widehat{P}_k = \sqrt{\frac{2k+1}{2}}\, P_k \qquad \text{and} \qquad \langle\, f, \widehat{P}_k\,\rangle = \int_{-1}^{1} f(x)\, \widehat{P}_k(x)\, dx.$$

The partial sum operator $S_n(f) = \sum_{k=0}^n \langle\, f, \widehat{P}_k\,\rangle\, \widehat{P}_k$ is a linear projection onto $\mathcal{P}_n$ and may be written as

$$S_n(f)(x) = \int_{-1}^{1} f(t)\, K_n(t, x)\, dt,$$

where $K_n(t, x) = \sum_{k=0}^n \widehat{P}_k(t)\, \widehat{P}_k(x)$. (Why?)

Because the polynomials $\widehat{P}_k$ are orthonormal, we have

$$\sum_{k=0}^{n} |\langle\, f, \widehat{P}_k\,\rangle|^2 \;=\; \|S_n(f)\|_2^2 \;\le\; \|f\|_2^2 \;=\; \sum_{k=0}^{\infty} |\langle\, f, \widehat{P}_k\,\rangle|^2,$$

and so the generalized Fourier coefficients $\langle\, f, \widehat{P}_k\,\rangle$ are square summable; in particular, $\langle\, f, \widehat{P}_k\,\rangle \to 0$ as $k \to \infty$. As in the case of Fourier series, the fact that the polynomials (i.e., the span of the $\widehat{P}_k$) are dense in $C[\,a, b\,]$ implies that $S_n(f)$ actually converges to f in the $\|\cdot\|_2$ norm. These same observations remain valid for any sequence of orthogonal polynomials. The real question remains, just as with Fourier series, whether $S_n(f)$ is a good uniform (or even pointwise) approximation to f.

If you're willing to swallow the fact that $|P_n(x)| \le 1$, we get

$$|K_n(t, x)| \;\le\; \sum_{k=0}^{n} \sqrt{\frac{2k+1}{2}}\, \sqrt{\frac{2k+1}{2}} \;=\; \frac{1}{2}\sum_{k=0}^{n}(2k+1) \;=\; \frac{(n+1)^2}{2}.$$

Hence, $\|S_n(f)\| \le (n+1)^2\|f\|$. That is, the Lebesgue numbers for this process are at most $(n+1)^2$. The analogue of Lebesgue's theorem in this case would then read:

$$\|f - S_n(f)\| \;\le\; Cn^2 E_n(f).$$

Thus, $S_n(f) \rightrightarrows f$ whenever $n^2 E_n(f) \to 0$, and Jackson's theorem tells us when this will happen: *If f is twice continuously differentiable, then the Fourier-Legendre series for f converges uniformly to f on $[-1,1]$.*

The Christoffel-Darboux Identity

It would also be of interest to have a closed form for $K_n(t,x)$. That this is indeed always possible, for any sequence of orthogonal polynomials, is a very important fact.

Using our original notation, let (Q_n) be the sequence of monic orthogonal polynomials corresponding to a given weight w, and let $(\widehat{Q}_n)$ be the *orthonormal* counterpart of (Q_n); in other words, $Q_n = \lambda_n \widehat{Q}_n$, where $\lambda_n = \sqrt{\langle Q_n, Q_n \rangle}$. It will help things here if you recall (from Remarks 8.2 (1)) that $\lambda_n^2 = b_n \lambda_{n-1}^2$.

As with the Legendre polynomials, each $f \in C[\,a,b\,]$ is represented by a generalized Fourier series $\sum_k \langle f, \widehat{Q}_k \rangle \widehat{Q}_k$ with partial sum operator

$$S_n(f)(x) = \int_a^b f(t)\, K_n(t,x)\, w(t)\, dt,$$

where $K_n(t,x) = \sum_{k=0}^n \widehat{Q}_k(t)\, \widehat{Q}_k(x)$. As before, S_n is a projection onto $\mathcal{P}_n$; in particular, $S_n(1) = 1$ for every n.

Theorem 8.12. (Christoffel-Darboux) *The kernel $K_n(t,x)$ can be written*

$$\sum_{k=0}^n \widehat{Q}_k(t)\, \widehat{Q}_k(x) \;=\; \lambda_{n+1}\lambda_n^{-1}\, \frac{\widehat{Q}_{n+1}(t)\, \widehat{Q}_n(x) - \widehat{Q}_n(t)\, \widehat{Q}_{n+1}(x)}{t-x}.$$

Proof. We begin with the standard recurrence formulas

$$\begin{aligned} Q_{n+1}(t) &= (t-a_n)\, Q_n(t) - b_n Q_{n-1}(t) \\ Q_{n+1}(x) &= (x-a_n)\, Q_n(x) - b_n Q_{n-1}(x) \end{aligned}$$

(where $b_0 = 0$). Multiplying the first by $Q_n(x)$, the second by $Q_n(t)$, and subtracting:

$$\begin{aligned} &Q_{n+1}(t)\, Q_n(x) - Q_n(t)\, Q_{n+1}(x) \\ &\quad = (t-x)\, Q_n(t)\, Q_n(x) \;+\; b_n \big[\, Q_n(t)\, Q_{n-1}(x) \;-\; Q_n(x)\, Q_{n-1}(t)\,\big] \end{aligned}$$

(and again, $b_0 = 0$). If we divide both sides of this equation by λ_n^2 we get

$$\begin{aligned} &\lambda_n^{-2} \big[\, Q_{n+1}(t)\, Q_n(x) - Q_n(t)\, Q_{n+1}(x)\,\big] \\ &\quad = (t-x)\, \widehat{Q}_n(t)\, \widehat{Q}_n(x) \;+\; \lambda_{n-1}^{-2} \big[\, Q_n(t)\, Q_{n-1}(x) \;-\; Q_n(x)\, Q_{n-1}(t)\,\big]. \end{aligned}$$

Thus, we may repeat the process; arriving finally at

$$\lambda_n^{-2} \big[\, Q_{n+1}(t)\, Q_n(x) - Q_n(t)\, Q_{n+1}(x)\,\big] \;=\; (t-x) \sum_{k=0}^n \widehat{Q}_k(t)\, \widehat{Q}_k(x).$$

The Christoffel-Darboux identity now follows by writing $Q_n = \lambda_n \widehat{Q}_n$, etc. □

And we now have a version of the Dini-Lipschitz theorem (Theorem 7.3).

Theorem 8.13. *Let $f \in C[\,a,b\,]$ and suppose that at some point x_0 in $[\,a,b\,]$ we have*

(i) *f is Lipschitz at x_0; that is, $|f(x_0) - f(x)| \le K|x_0 - x|$ for some constant K and all x in $[\,a,b\,]$; and*

(ii) *the sequence $\big(\widehat{Q}_n(x_0)\big)$ is bounded.*

Then, the series $\sum_k \langle\, f, \widehat{Q}_k\,\rangle\, \widehat{Q}_k(x_0)$ converges to $f(x_0)$.

Proof. First note that the sequence $\lambda_{n+1}\lambda_n^{-1}$ is bounded: Indeed, by Cauchy-Schwarz,

$$\begin{aligned}\lambda_{n+1}^2 &= \langle\, Q_{n+1}, Q_{n+1}\,\rangle \;=\; \langle\, Q_{n+1}, x\,Q_n\,\rangle \\ &\le \|Q_{n+1}\|_2 \cdot \|\,x\,\| \cdot \|Q_n\|_2 \;=\; \max\{|a|,|b|\}\,\lambda_{n+1}\lambda_n.\end{aligned}$$

Thus, $\lambda_{n+1}\lambda_n^{-1} \le c = \max\{|a|,|b|\}$. Now, using the Christoffel-Darboux identity,

$$\begin{aligned}S_n(f)(x_0) - f(x_0) &= \textstyle\int_a^b \big[\, f(t) - f(x_0)\,\big]\, K_n(t,x_0)\, w(t)\, dt \\ &= \lambda_{n+1}\lambda_n^{-1} \int_a^b \frac{f(t)-f(x_0)}{t-x_0}\,\big[\,\widehat{Q}_{n+1}(t)\,\widehat{Q}_n(x_0) - \widehat{Q}_n(t)\,\widehat{Q}_{n+1}(x_0)\,\big]\, w(t)\, dt \\ &= \lambda_{n+1}\lambda_n^{-1}\,\big[\,\langle\, h, \widehat{Q}_{n+1}\,\rangle\, \widehat{Q}_n(x_0) - \langle\, h, \widehat{Q}_n\,\rangle\, \widehat{Q}_{n+1}(x_0)\,\big],\end{aligned}$$

where $h(t) = (f(t) - f(x_0))/(t - x_0)$. But h is bounded (and continuous everywhere except, possibly, at x_0) by hypothesis (i), $\lambda_{n+1}\lambda_n^{-1}$ is bounded, and $\widehat{Q}_n(x_0)$ is bounded by hypothesis (ii). All that remains is to notice that the numbers $\langle\, h, \widehat{Q}_n\,\rangle$ are the generalized Fourier coefficients of the bounded, Riemann integrable function h, and so must tend to zero (because, in fact, they're even square summable). □

We end this chapter with a negative result, due to Nikolaev:

Theorem 8.14. *There is no weight w such that every $f \in C[\,a,b\,]$ has a uniformly convergent expansion in terms of orthogonal polynomials. In fact, given any w, there is always some f for which $\|f - S_n(f)\|$ is unbounded.*

Problems

▷ 1. Prove that every inner product norm is strictly convex. Specifically, let $\langle \cdot, \cdot \rangle$ be an inner product on a vector space X, and let $\|x\| = \sqrt{\langle x, x \rangle}$ be the associated norm. Show that:

(a) $\|x+y\|^2 + \|x-y\|^2 = 2\,(\|x\|^2 + \|y\|^2)$ for all $x, y \in X$ (the parallelogram identity).

(b) If $\|x\| = r = \|y\|$ and if $\|x - y\| = \delta$, then $\left\| \frac{x+y}{2} \right\|^2 = r^2 - (\delta/2)^2$. In particular, $\left\| \frac{x+y}{2} \right\| < r$ whenever $x \neq y$.

The remaining problems follow the notation given on page 79.

▷ 2. (a) Show that the expression $\|f\|_1 = \int_a^b |f(t)| w(t)\, dt$ also defines a norm on $C[\,a, b\,]$.

(b) Given any f in $C[\,a, b\,]$, show that $\|f\|_1 \le c\|f\|_2$ and $\|f\|_2 \le c\|f\|$, where $c = \left(\int_a^b w(t)\, dt \right)^{1/2}$.

(c) Conclude that the polynomials are dense in $C[\,a, b\,]$ under all three of the norms $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|$.

(d) Show that $C[\,a, b\,]$ is *not* complete under either of the norms $\|\cdot\|_1$ or $\|\cdot\|_2$.

3. Check that Q_n is a monic polynomial of degree exactly n.

4. If (P_n) is another sequence of orthogonal polynomials such that P_n has degree exactly n, for each n, show that $P_n = \alpha_n Q_n$ for some $\alpha_n \neq 0$. In particular, if P_n is a monic polynomial, then $P_n = Q_n$. [Hint: Choose α_n so that $P_n - \alpha_n Q_n \in \mathcal{P}_{n-1}$ and note that $(P_n - \alpha_n Q_n) \perp \mathcal{P}_{n-1}$. Conclude that $P_n - \alpha_n Q_n = 0$.]

5. Given $w > 0$, $f \in C[\,a, b\,]$, and $n \ge 1$, show that $p^* \in \mathcal{P}_{n-1}$ is the least-squares approximation to f out of $\mathcal{P}_{n-1}$ (with respect to w) if and only if $\langle f - p^*, p \rangle = 0$ for every $p \in \mathcal{P}_{n-1}$; that is, if and only if $(f - p^*) \perp \mathcal{P}_{n-1}$.

6. In the notation of Problem 5, show that $f - p^*$ has at least n distinct zeros in (a, b).

7. If $w > 0$, show that the least-squares approximation to $f(x) = x^n$ out of $\mathcal{P}_{n-1}$ (relative to w) is $q_{n-1}^*(x) = x^n - Q_n(x)$.

▷ 8. Given $f \in C[\,a, b\,]$, let p_n^* denote the best *uniform* approximation to f out of $\mathcal{P}_n$ and let q_n^* denote the least-squares approximation to f out of $\mathcal{P}_n$. Show that $\|f - q_n^*\|_2 \le \|f - p_n^*\|_2$ and conclude that $\|f - q_n^*\|_2 \to 0$ as $n \to \infty$.

9. Show that the Chebyshev polynomials of the first kind, (T_n), and of the second kind, (U_n), satisfy the identities

$$T_n(x) = U_n(x) - x\, U_{n-1}(x)$$

and

$$(1 - x^2)\, U_{n-1}(x) = x\, T_n(x) - T_{n+1}(x).$$

10. Show that the Chebyshev polynomials of the second kind, (U_n), satisfy the recurrence relation

$$U_{n+1}(x) = 2x\, U_n(x) - U_{n-1}(x), \quad n \ge 1,$$

where $U_0(x) = 1$ and $U_1(x) = 2x$. [Compare this with the recurrence relation satisfied by the T_n.]