Chapter 4

Orthogonality and Least-Squares Approximation

4.1 Introduction — from minimax to least squares

The Chebyshev polynomials have been shown in Chapter 3 to be unique among all polynomials in possessing a minimax property (Corollaries 3.4B, 3.5A), earning them a central role in the study of uniform (or $\mathcal{L}_\infty$) approximation. This property is remarkable enough, but the four families of Chebyshev polynomials have a second and equally important property, in that each is a family of orthogonal polynomials. Thus, the Chebyshev polynomials have an important role in $\mathcal{L}_2$ or least-squares approximation, too. This link with $\mathcal{L}_2$ approximation is important in itself but, in addition, it enables ideas of orthogonality to be exploited in such areas as Chebyshev series expansions and Galerkin methods for differential equations.

Orthogonal polynomials have a great variety and wealth of properties, many of which are noted in this chapter. Indeed, some of these properties take a very concise form in the case of the Chebyshev polynomials, making Chebyshev polynomials of leading importance among orthogonal polynomials — second perhaps to Legendre polynomials (which have a unit weight function), but having the advantage over the Legendre polynomials that the locations of their zeros are known analytically. Moreover, along with the Legendre polynomials, the Chebyshev polynomials belong to an exclusive band of orthogonal polynomials, known as *Jacobi polynomials*, which correspond to weight functions of the form $(1-x)^\alpha(1+x)^\beta$ and which are solutions of Sturm–Liouville equations.

The Chebyshev polynomials have further properties, which are peculiar to them and have a trigonometric origin, namely various kinds of discrete orthogonality over the zeros of Chebyshev polynomials of higher degree. In consequence, interpolation at Chebyshev zeros can be achieved exceptionally inexpensively (Chapter 6) and Gauss quadrature methods based on Chebyshev zeros are extremely convenient (Chapter 8).

The continuous and discrete orthogonality of the Chebyshev polynomials may be viewed as a direct consequence of the orthogonality of sine and cosine functions of multiple angles, a central feature in the study of Fourier series. It is likely, therefore, that a great deal may be learned about Chebyshev series by studying their links with Fourier series (or, in the complex plane, Laurent series); this is considered in Chapter 5.

Finally, the Chebyshev polynomials are orthogonal not only as polynomials in the real variable x on the real interval $[-1, 1]$ but also as polynomials in a complex variable z on elliptical contours and domains of the complex plane (the foci of the ellipses being at -1 and $+1$). This property is exploited in fields such as crack problems in fracture mechanics (Gladwell & England 1977) and two-dimensional aerodynamics (Fromme & Golberg 1979, Fromme & Golberg 1981), which rely on complex-variable techniques. More generally, however, many real functions may be extended into analytic functions, and Chebyshev polynomials are remarkably robust in approximating on $[-1, 1]$ functions which have complex poles close to that interval. This is a consequence of the fact that the interval $[-1, 1]$ may be enclosed in an arbitrarily thin ellipse which excludes nearby singularities.

4.2 Orthogonality of Chebyshev polynomials

4.2.1 Orthogonal polynomials and weight functions

Definition 4.1 *Two functions $f(x)$ and $g(x)$ in $\mathcal{L}_2[a, b]$ are said to be* orthogonal *on the interval $[a, b]$ with respect to a given continuous and non-negative weight function $w(x)$ if*

$$\int_a^b w(x)f(x)g(x)\,\mathrm{d}x = 0. \tag{4.1}$$

If, for convenience, we use the 'inner product' notation

$$\langle f\,,\,g\rangle = \int_a^b w(x)f(x)g(x)\,\mathrm{d}x, \tag{4.2}$$

where w, f and g are functions of x on $[a, b]$, then the orthogonality condition (4.1) is equivalent to saying that f is orthogonal to g if

$$\langle f\,,\,g\rangle = 0. \tag{4.3}$$

The formal definition of an inner product (in the context of real functions of a real variable — see Definition 4.3 for the complex case) is as follows:

Definition 4.2 *An* inner product *$\langle \cdot\,,\,\cdot\rangle$ is a bilinear function of elements $f, g, h, \ldots$ of a vector space that satisfies the axioms:*

1. *$\langle f\,,\,f\rangle \geq 0$ with equality if and only if $f \equiv 0$;*
2. *$\langle f\,,\,g\rangle = \langle g\,,\,f\rangle$;*
3. *$\langle f + g\,,\,h\rangle = \langle f\,,\,h\rangle + \langle g\,,\,h\rangle$;*
4. *$\langle \alpha f\,,\,g\rangle = \alpha\,\langle f\,,\,g\rangle$ for any scalar α.*

An inner product defines an $\mathcal{L}_2$-type norm

$$\|f\| = \|f\|_2 := \sqrt{\langle f\,,\, f\rangle}. \tag{4.4}$$

We shall adopt the inner product (4.2) (with various weight functions) and the associated $\mathcal{L}_2$ norm (4.4), which is identical to that defined in Chapter 3 (3.4), through most of the remainder of this chapter.

Here we shall in particular be concerned with families of orthogonal polynomials $\{\phi_i(x),\ i = 0, 1, 2, \ldots\}$ where ϕ_i is of degree i exactly, defined so that

$$\langle \phi_i\,,\, \phi_j\rangle = 0 \quad (i \neq j). \tag{4.5}$$

Clearly, since $w(x)$ is non-negative,

$$\langle \phi_i\,,\, \phi_i\rangle = \|\phi_i\|^2 > 0. \tag{4.6}$$

The requirement that ϕ_i should be of exact degree i, together with the orthogonality condition (4.5), defines each polynomial ϕ_i uniquely apart from a multiplicative constant (see Problem 3). The definition may be made unique by fixing the value of $\langle \phi_i\,,\, \phi_i\rangle$ or of its square root $\|\phi_i\|$. In particular, we say that the family is *orthonormal* if, in addition to (4.5), the functions $\{\phi_i(x)\}$ satisfy

$$\|\phi_i\| = 1 \text{ for all } i. \tag{4.7}$$

4.2.2 Chebyshev polynomials as orthogonal polynomials

If we define the inner product (4.2) using the interval and weight function

$$[a, b] = [-1, 1], \quad w(x) = (1 - x^2)^{-\frac{1}{2}}, \tag{4.8}$$

then we find that the first kind Chebyshev polynomials satisfy

$$\begin{aligned} \langle T_i\,,\, T_j\rangle &= \int_{-1}^{1} \frac{T_i(x)T_j(x)}{\sqrt{1-x^2}}\,\mathrm{d}x \\ &= \int_0^{\pi} \cos i\theta\, \cos j\theta\, \mathrm{d}\theta \end{aligned} \tag{4.9}$$

(shown by setting $x = \cos\theta$ and using the relations $T_i(x) = \cos i\theta$ and $\mathrm{d}x = -\sin\theta\,\mathrm{d}\theta = -\sqrt{1-x^2}\,\mathrm{d}\theta$).

Now, for $i \neq j$,

$$\begin{aligned} \int_0^{\pi} \cos i\theta\, \cos j\theta\, \mathrm{d}\theta &= \tfrac{1}{2}\int_0^{\pi} [\cos(i+j)\theta + \cos(i-j)\theta]\,\mathrm{d}\theta \\ &= \tfrac{1}{2}\left[\frac{\sin(i+j)\theta}{i+j} + \frac{\sin(i-j)\theta}{i-j}\right]_0^{\pi} = 0. \end{aligned}$$

Hence

$$\langle T_i\,,\,T_j\rangle = 0 \quad (i \neq j), \tag{4.10}$$

and $\{T_i(x),\ i = 0, 1, \ldots\}$ forms an orthogonal polynomial system on $[-1, 1]$ with respect to the weight $(1 - x^2)^{-\frac{1}{2}}$.

The norm of T_i is given by

$$\begin{aligned} \|T_i\|^2 &= \langle T_i\,,\,T_i\rangle \\ &= \int_0^\pi (\cos i\theta)^2\,\mathrm{d}\theta \\ &= \tfrac{1}{2}\int_0^\pi (1 + \cos 2i\theta)\,\mathrm{d}\theta \\ &= \tfrac{1}{2}\left[\theta + \frac{\sin 2i\theta}{2i}\right]_0^\pi \quad (i \neq 0) \\ &= \tfrac{1}{2}\pi, \end{aligned} \tag{4.11a}$$

while

$$\|T_0\|^2 = \langle T_0\,,\,T_0\rangle = \langle 1\,,\,1\rangle = \pi. \tag{4.11b}$$

The system $\{T_i\}$ is therefore not orthonormal. We could, if we wished, scale the polynomials to derive the orthonormal system

$$\sqrt{1/\pi}\,T_0(x), \left\{\sqrt{2/\pi}\,T_i(x),\ i = 1, 2, \ldots\right\},$$

but the resulting irrational coefficients usually make this inconvenient. It is simpler in practice to adopt the $\{T_i\}$ we defined initially, taking note of the values of their norms (4.11).

The second, third and fourth kind Chebyshev polynomials are also orthogonal systems on $[-1, 1]$, with respect to appropriate weight functions:

- $U_i(x)$ are orthogonal with respect to $w(x) = (1 - x^2)^{\frac{1}{2}}$;
- $V_i(x)$ are orthogonal with respect to $w(x) = (1 + x)^{\frac{1}{2}}(1 - x)^{-\frac{1}{2}}$;
- $W_i(x)$ are orthogonal with respect to $w(x) = (1 + x)^{-\frac{1}{2}}(1 - x)^{\frac{1}{2}}$.

These results are obtained from trigonometric relations as follows (using the appropriate definition of $\langle\cdot\,,\,\cdot\rangle$ in each case):

$$\begin{aligned} \langle U_i\,,\,U_j\rangle &= \int_{-1}^1 (1 - x^2)^{\frac{1}{2}}\,U_i(x)U_j(x)\,\mathrm{d}x \\ &= \int_{-1}^1 (1 - x^2)^{-\frac{1}{2}}\,(1 - x^2)^{\frac{1}{2}}U_i(x)\,(1 - x^2)^{\frac{1}{2}}U_j(x)\,\mathrm{d}x \end{aligned}$$

$$
\begin{aligned}
&= \int_0^\pi \sin(i+1)\theta\ \sin(j+1)\theta\,\mathrm{d}\theta \\
&\qquad (\text{since } \sin\theta\, U_i(x) = \sin(i+1)\theta) \\
&= \tfrac{1}{2}\int_0^\pi [\cos(i-j)\theta - \cos(i+j+2)\theta]\,\mathrm{d}\theta \\
&= 0 \quad (i \neq j).
\end{aligned}
$$

$$
\begin{aligned}
\langle V_i\,,\, V_j\rangle &= \int_{-1}^{1} (1-x)^{-\frac{1}{2}}(1+x)^{\frac{1}{2}}\ V_i(x)V_j(x)\,\mathrm{d}x \\
&= \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}}\ (1+x)^{\frac{1}{2}} V_i(x)\ (1+x)^{\frac{1}{2}} V_j(x)\,\mathrm{d}x \\
&= 2\int_0^\pi \cos(i+\tfrac{1}{2})\theta\ \cos(j+\tfrac{1}{2})\theta\,\mathrm{d}\theta \\
&\qquad (\text{since } (1+x)^{\frac{1}{2}} = (1+\cos\theta)^{\frac{1}{2}} = (2\cos^2\tfrac{1}{2}\theta)^{\frac{1}{2}} = \sqrt{2}\cos\tfrac{1}{2}\theta \\
&\qquad\quad \text{and } (1+x)^{\frac{1}{2}} V_i(x) = \sqrt{2}\cos(i+\tfrac{1}{2})\theta) \\
&= \int_0^\pi [\cos(i+j+1)\theta + \cos(i-j)\theta]\,\mathrm{d}\theta \\
&= 0 \quad (i \neq j).
\end{aligned}
$$

$$
\begin{aligned}
\langle W_i\,,\, W_j\rangle &= \int_{-1}^{1} (1+x)^{-\frac{1}{2}}(1-x)^{\frac{1}{2}}\ W_i(x)W_j(x)\,\mathrm{d}x \\
&= \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}}\ (1-x)^{\frac{1}{2}} W_i(x)\ (1-x)^{\frac{1}{2}} W_j(x)\,\mathrm{d}x \\
&= 2\int_0^\pi \sin(i+\tfrac{1}{2})\theta\ \sin(j+\tfrac{1}{2})\theta\,\mathrm{d}\theta \\
&\qquad (\text{since } (1-x)^{\frac{1}{2}} = (1-\cos\theta)^{\frac{1}{2}} = (2\sin^2\tfrac{1}{2}\theta)^{\frac{1}{2}} = \sqrt{2}\sin\tfrac{1}{2}\theta \\
&\qquad\quad \text{and } (1-x)^{\frac{1}{2}} W_i(x) = \sqrt{2}\sin(i+\tfrac{1}{2})\theta) \\
&= \int_0^\pi [\cos(i-j)\theta - \cos(i+j+1)\theta]\,\mathrm{d}\theta \\
&= 0 \quad (i \neq j).
\end{aligned}
$$

The normalisations that correspond to these polynomials are as follows (for all $i \geq 0$):

$$
\langle U_i\,,\, U_i\rangle = \|U_i\|^2 = \int_0^\pi \sin^2(i+1)\theta\,\mathrm{d}\theta = \tfrac{1}{2}\pi; \tag{4.12}
$$

$$\langle V_i\,,\,V_i\rangle = \|V_i\|^2 = 2\int_0^\pi \cos^2(i+\tfrac{1}{2})\theta\,\mathrm{d}\theta = \pi; \tag{4.13}$$

$$\langle W_i\,,\,W_i\rangle = \|W_i\|^2 = 2\int_0^\pi \sin^2(i+\tfrac{1}{2})\theta\,\mathrm{d}\theta = \pi. \tag{4.14}$$

(Remember that each of these three identities uses a different definition of the inner product $\langle\cdot\,,\,\cdot\rangle$, since the weights $w(x)$ differ.)

4.3 Orthogonal polynomials and best $\mathcal{L}_2$ approximations

In Chapter 3, we characterised a best $\mathcal{L}_\infty$ (minimax) polynomial approximation, by way of Chebyshev's theorem, and this led us to an equioscillation property. Now we consider the best $\mathcal{L}_2$ polynomial approximation of a given degree, which leads us to an orthogonality property.

The theorems in this section are valid not only for the inner product (4.2), but for any inner product $\langle\cdot\,,\,\cdot\rangle$ as defined by Definition 4.2.

Theorem 4.1 *The best $\mathcal{L}_2$ polynomial approximation $p_n^B(x)$ of degree n (or less) to a given ($\mathcal{L}_2$-integrable) function $f(x)$ is unique and is characterised by the (necessary and sufficient) property that*

$$\left\langle f - p_n^B\,,\,p_n\right\rangle = 0 \tag{4.15}$$

for any other polynomial p_n of degree n.

Proof: Write

$$e_n^B := f - p_n^B.$$

1. (Necessity) Suppose that, for some polynomial p_n,

$$\left\langle e_n^B\,,\,p_n\right\rangle \neq 0.$$

Then, for any real scalar multiplier λ,

$$\begin{aligned}
\left\|f - (p_n^B + \lambda p_n)\right\|^2 &= \left\|e_n^B - \lambda p_n\right\|^2 \\
&= \left\langle e_n^B - \lambda p_n\,,\,e_n^B - \lambda p_n\right\rangle \\
&= \left\langle e_n^B\,,\,e_n^B\right\rangle - 2\lambda\left\langle e_n^B\,,\,p_n\right\rangle + \lambda^2\langle p_n\,,\,p_n\rangle \\
&= \left\|e_n^B\right\|^2 - 2\lambda\left\langle e_n^B\,,\,p_n\right\rangle + \lambda^2\|p_n\|^2 \\
&< \left\|e_n^B\right\|^2 \quad \text{for some small } \lambda \text{ of the same sign as } \left\langle e_n^B\,,\,p_n\right\rangle.
\end{aligned}$$

Hence $p_n^B + \lambda p_n$ is a better approximation than p_n^B for this value of λ, contradicting the assertion that p_n^B is a best approximation.

2. **(Sufficiency)** Suppose that (4.15) holds and that q_n is any specified polynomial of degree n, not identical to p_n^B. Then

$$\begin{aligned}
\|f - q_n\|^2 - \left\|f - p_n^B\right\|^2 &= \left\|e_n^B + (p_n^B - q_n)\right\|^2 - \left\|e_n^B\right\|^2 \\
&= \left\langle e_n^B + (p_n^B - q_n)\,,\, e_n^B + (p_n^B - q_n)\right\rangle - \left\langle e_n^B\,,\, e_n^B\right\rangle \\
&= \left\langle p_n^B - q_n\,,\, p_n^B - q_n\right\rangle + 2\left\langle e_n^B\,,\, p_n^B - q_n\right\rangle \\
&= \left\|p_n^B - q_n\right\|^2 + 0, \text{ from (4.15)} \\
&> 0.
\end{aligned}$$

Therefore $\|f - q_n\|^2 > \|f - p_n^B\|^2$.

Since q_n is arbitrary, p_n^B must be a best $\mathcal{L}_2$ approximation. It must also be unique, since otherwise we could have taken q_n to be another best approximation and obtained the last inequality as a contradiction. ●●

Corollary 4.1A *If $\{\phi_n\}$ (ϕ_i being of exact degree i) is an orthogonal polynomial system on $[a, b]$, then:*

1. *the zero function is the best $\mathcal{L}_2$ polynomial approximation of degree $(n-1)$ to ϕ_n on $[a, b]$;*

2. *ϕ_n is the best $\mathcal{L}_2$ approximation to zero on $[a, b]$ among polynomials of degree n with the same leading coefficient.*

Proof:

1. Any polynomial p_{n-1} of degree $n-1$ can be written in the form

$$p_{n-1} = \sum_{i=0}^{n-1} c_i \phi_i.$$

Then

$$\begin{aligned}
\langle \phi_n - 0\,,\, p_{n-1}\rangle &= \left\langle \phi_n\,,\, \sum_{i=0}^{n-1} c_i\phi_i\right\rangle \\
&= \sum_{i=0}^{n-1} c_i \langle \phi_n\,,\, \phi_i\rangle \\
&= 0 \text{ by the orthogonality of } \{\phi_i\}.
\end{aligned}$$

The result follows from Theorem 4.1.

2. Let q_n be any other polynomial of degree n having the same leading coefficient as ϕ_n. Then $q_n - \phi_n$ is a polynomial of degree $n-1$. We can therefore write

$$q_n - \phi_n = \sum_{i=0}^{n-1} c_i \phi_i$$

and deduce from the orthogonality of $\{\phi_i\}$ that

$$\langle \phi_n \,,\, q_n - \phi_n \rangle = 0. \tag{4.16}$$

Now we have

$$\begin{aligned} \|q_n\|^2 - \|\phi_n\|^2 &= \langle q_n \,,\, q_n \rangle - \langle \phi_n \,,\, \phi_n \rangle \\ &= \langle q_n - \phi_n \,,\, q_n - \phi_n \rangle - 2\,\langle \phi_n \,,\, q_n - \phi_n \rangle \\ &= \|q_n - \phi_n\|^2 \,, \text{ using (4.16)} \\ &> 0. \end{aligned}$$

Therefore ϕ_n is the best approximation to zero. ●●

The interesting observation that follows from Corollary 4.1A is that every polynomial in an orthogonal system has a minimal $\mathcal{L}_2$ property — analogous to the minimax property of the Chebyshev polynomials. Indeed, the four kinds of Chebyshev polynomials T_n, U_n, V_n, W_n, being orthogonal polynomials, each have a minimal property on $[-1, 1]$ with respect to their respective weight functions

$$\frac{1}{\sqrt{1-x^2}}, \; \sqrt{1-x^2}, \; \sqrt{\frac{1+x}{1-x}}, \; \sqrt{\frac{1-x}{1+x}}$$

over all polynomials with the same leading coefficients.

The main result above, namely Theorem 4.1, is essentially a generalisation of the statement that *the shortest distance from a point to a plane is in the direction of a vector perpendicular to all vectors in that plane.*

Theorem 4.1 is important in that it leads to a very direct algorithm for determining the best $\mathcal{L}_2$ polynomial approximation p_n^B to f:

Corollary 4.1B *The best $\mathcal{L}_2$ polynomial approximation p_n^B of degree n to f may be expressed in terms of the orthogonal polynomial family $\{\phi_i\}$ in the form*

$$p_n^B = \sum_{i=0}^{n} c_i \phi_i,$$

where

$$c_i = \frac{\langle f \,,\, \phi_i \rangle}{\langle \phi_i \,,\, \phi_i \rangle}.$$

Proof: For $k = 0, 1, \ldots, n$

$$\begin{aligned}\left\langle f - p_n^B \, , \, \phi_k \right\rangle &= \left\langle f - \sum_{i=0}^{n} c_i \phi_i \, , \, \phi_k \right\rangle \\ &= \langle f \, , \, \phi_k \rangle - \sum_{i=0}^{n} c_i \, \langle \phi_i \, , \, \phi_k \rangle \\ &= \langle f \, , \, \phi_k \rangle - c_k \, \langle \phi_k \, , \, \phi_k \rangle \\ &= 0, \text{ by definition of } c_k. \end{aligned} \tag{4.17}$$

Now, any polynomial p_n can be written as

$$p_n = \sum_{i=0}^{n} d_i \phi_i,$$

and hence

$$\begin{aligned}\left\langle f - p_n^B \, , \, p_n \right\rangle &= \sum_{i=0}^{n} d_i \left\langle f - p_n^B \, , \, \phi_i \right\rangle \\ &= 0 \text{ by (4.17).}\end{aligned}$$

Thus p_n^B is the best approximation by Theorem 4.1. ●●

Example 4.1: To illustrate Corollary 4.1B, suppose that we wish to determine the best $\mathcal{L}_2$ linear approximation p_1^B to $f(x) = 1 - x^2$ on $[-1, 1]$, with respect to the weight $w(x) = (1 - x^2)^{-\frac{1}{2}}$. In this case $\{T_i(x)\}$ is the appropriate orthogonal system and hence

$$p_1^B = c_0 T_0(x) + c_1 T_1(x)$$

where, by (4.17),

$$\begin{aligned} c_0 &= \frac{\langle f \, , \, T_0 \rangle}{\langle T_0 \, , \, T_0 \rangle} = \frac{\int_{-1}^{1} (1 - x^2)^{-\frac{1}{2}} (1 - x^2) \, \mathrm{d}x}{\pi}, \\ c_1 &= \frac{\langle f \, , \, T_0 \rangle}{\langle T_0 \, , \, T_0 \rangle} = \frac{\int_{-1}^{1} (1 - x^2)^{-\frac{1}{2}} (1 - x^2) x \, \mathrm{d}x}{\frac{1}{2}\pi}. \end{aligned}$$

Substituting $x = \cos\theta$,

$$\begin{aligned} c_0 &= \frac{1}{\pi} \int_0^{\pi} \sin^2\theta \, \mathrm{d}\theta = \frac{1}{2\pi} \int_0^{\pi} (1 - \cos 2\theta) \, \mathrm{d}\theta = \tfrac{1}{2}, \\ c_1 &= \frac{2}{\pi} \int_0^{\pi} \sin^2\theta \, \cos\theta \, \mathrm{d}\theta = \frac{2}{\pi} \left[\tfrac{1}{3} \sin^3\theta \right]_0^{\pi} = 0 \end{aligned}$$

and therefore

$$p_1^B = \tfrac{1}{2} T_0(x) + 0 \, T_1(x) = \tfrac{1}{2},$$

so that the linear approximation reduces to a constant in this case.

4.3.1 Orthogonal polynomial expansions

On the assumption that it is possible to expand a given function $f(x)$ in a (suitably convergent) series based on a system $\{\phi_k\}$ of polynomials orthogonal over the interval $[a, b]$, ϕ_k being of exact degree k, we may write

$$f(x) = \sum_{k=0}^{\infty} c_i \phi_i(x), \quad x \in [a, b]. \tag{4.18}$$

It follows, by taking inner products with ϕ_k, that

$$\langle f \, , \, \phi_k \rangle = \sum_{k=0}^{\infty} c_i \langle \phi_i \, , \, \phi_k \rangle = c_k \langle \phi_k \, , \, \phi_k \rangle \, ,$$

since $\langle \phi_i \, , \, \phi_k \rangle = 0$ for $i \neq k$. This is identical to the formula for c_k given in Corollary 4.1B. Thus (applying the same corollary) an orthogonal expansion has the property that its partial sum of degree n is the best $\mathcal{L}_2$ approximation of degree n to its infinite sum. Hence it is an ideal expansion to use in the $\mathcal{L}_2$ context. In particular, the four Chebyshev series expansions have this property on $[-1, 1]$ with respect to their respective weight functions $(1 + x)^{\pm\frac{1}{2}}(1 - x)^{\pm\frac{1}{2}}$.

We shall have much more to say on this topic in Chapter 5.

4.3.2 Convergence in $\mathcal{L}_2$ of orthogonal expansions

Convergence questions will be considered in detail in Chapter 5, where we shall restrict attention to Chebyshev polynomials and use Fourier series theory. However, we may easily make some deductions from general orthogonal polynomial properties.

In particular, if f is continuous, then we know (Theorem 3.2) that arbitrarily accurate polynomial approximations exist in $\mathcal{C}[a, b]$, and it follows from Lemma 3.1 that these are also arbitrarily accurate in $\mathcal{L}_2[a, b]$. However, we have shown in Section 4.3.1 that the nth degree polynomial, $P_n(x)$ say, obtained by truncating an orthogonal polynomial expansion is a best $\mathcal{L}_2$ approximation. Hence (*a fortiori*) P_n must also achieve an arbitrarily small $\mathcal{L}_2$ error $\|f - P_n\|_2$ for sufficiently large n. This gives the following result.

Theorem 4.2 *If f is in $\mathcal{C}[a, b]$, then its expansion in orthogonal polynomials converges in $\mathcal{L}_2$ (with respect to the appropriate weight function).*

In Chapter 5, we obtain much more powerful convergence results for Chebyshev series, ensuring $\mathcal{L}_2$ convergence of the series itself for f in $L_2[a, b]$ and L_∞ convergence of Cesàro sums of the series for f in $\mathcal{C}[a, b]$.

4.4 Recurrence relations

Using the inner product (4.2), namely

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

we note that

$$\langle f\,,\,g\rangle = \langle g\,,\,f\rangle\,, \tag{4.19}$$

$$\langle xf\,,\,g\rangle = \langle f\,,\,xg\rangle\,. \tag{4.20}$$

The following formulae uniquely define an orthogonal polynomial system $\{\phi_i\}$, in which ϕ_i is a monic polynomial (i.e., a polynomial with a leading coefficient of unity) of exact degree i.

Theorem 4.3 *The unique system of monic polynomials $\{\phi_i\}$, with ϕ_i of exact degree i, which are orthogonal on $[a,b]$ with respect to $w(x)$ are defined by*

$$\begin{aligned} \phi_0(x) &= 1, \\ \phi_1(x) &= x - a_1, \\ \phi_n(x) &= (x-a_n)\phi_{n-1}(x) - b_n\phi_{n-2}(x), \end{aligned} \tag{4.21}$$

where

$$a_n = \frac{\langle x\phi_{n-1}\,,\,\phi_{n-1}\rangle}{\langle \phi_{n-1}\,,\,\phi_{n-1}\rangle}, \qquad b_n = \frac{\langle \phi_{n-1}\,,\,\phi_{n-1}\rangle}{\langle \phi_{n-2}\,,\,\phi_{n-2}\rangle}.$$

Proof: This is readily shown by induction on n. It is easy to show that the polynomials ϕ_n generated by (4.21) are all monic. We assume that the polynomials ϕ_0, ϕ_1, ..., ϕ_{n-1} are orthogonal, and we then need to test that ϕ_n, as given by (4.21), is orthogonal to ϕ_k $(k = 0, 1, \ldots, n-1)$.

The polynomial $x\phi_k$ is a monic polynomial of degree $k+1$, expressible in the form

$$x\phi_k(x) = \phi_{k+1}(x) + \sum_{i=1}^{k} c_i\phi_i(x),$$

so that, using (4.20),

$$\begin{aligned} \langle x\phi_{n-1}\,,\,\phi_k\rangle &= \langle \phi_{n-1}\,,\,x\phi_k\rangle = 0 \ (k < n-2), \\ \langle x\phi_{n-1}\,,\,\phi_{n-2}\rangle &= \langle \phi_{n-1}\,,\,x\phi_{n-2}\rangle = \langle \phi_{n-1}\,,\,\phi_{n-1}\rangle\,. \end{aligned}$$

For $k < n-2$, then, we have

$$\langle \phi_n\,,\,\phi_k\rangle = \langle x\phi_{n-1}\,,\,\phi_k\rangle - a_n\langle \phi_{n-1}\,,\,\phi_k\rangle - b_n\langle \phi_{n-2}\,,\,\phi_k\rangle = 0,$$

while

$$\begin{aligned}
\langle \phi_n\,,\,\phi_{n-2}\rangle &= \langle x\phi_{n-1}\,,\,\phi_{n-2}\rangle - a_n\,\langle \phi_{n-1}\,,\,\phi_{n-2}\rangle - b_n\,\langle \phi_{n-2}\,,\,\phi_{n-2}\rangle \\
&= \langle \phi_{n-1}\,,\,\phi_{n-1}\rangle - 0 - \langle \phi_{n-1}\,,\,\phi_{n-1}\rangle = 0, \\
\langle \phi_n\,,\,\phi_{n-1}\rangle &= \langle x\phi_{n-1}\,,\,\phi_{n-1}\rangle - a_n\,\langle \phi_{n-1}\,,\,\phi_{n-1}\rangle - b_n\,\langle \phi_{n-2}\,,\,\phi_{n-1}\rangle \\
&= \langle x\phi_{n-1}\,,\,\phi_{n-1}\rangle - \langle x\phi_{n-1}\,,\,\phi_{n-1}\rangle - 0 = 0.
\end{aligned}$$

Starting the induction is easy, and the result follows. ●●

We have already established a recurrence relation for each of the four kinds of Chebyshev polynomials. We can verify that (4.21) leads to the same recurrences.

Consider the case of the polynomials of the first kind. We convert $T_n(x)$ to a monic polynomial by writing $\phi_0 = T_0$, $\phi_n = 2^{1-n}T_n$ $(n > 0)$. Then we can find the inner products:

$$\begin{aligned}
\langle T_0\,,\,T_0\rangle &= \int_{-1}^{1} \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x = \int_0^{\pi} \mathrm{d}\theta = \pi, \\
\langle xT_0\,,\,T_0\rangle &= \int_{-1}^{1} \frac{x}{\sqrt{1-x^2}}\,\mathrm{d}x = \int_0^{\pi} \cos\theta\,\mathrm{d}\theta = 0, \\
\langle T_n\,,\,T_n\rangle &= \int_{-1}^{1} \frac{T_n(x)^2}{\sqrt{1-x^2}}\,\mathrm{d}x = \int_0^{\pi} \cos^2 n\theta\,\mathrm{d}\theta = \tfrac{1}{2}\pi, \\
\langle xT_n\,,\,T_n\rangle &= \int_{-1}^{1} \frac{xT_n(x)^2}{\sqrt{1-x^2}}\,\mathrm{d}x = \int_0^{\pi} \cos\theta\,\cos^2 n\theta\,\mathrm{d}\theta = 0.
\end{aligned}$$

Therefore $a_1 = 0$, $a_n = 0$ $(n > 1)$, and

$$\begin{aligned}
b_2 &= \frac{\langle \phi_1\,,\,\phi_1\rangle}{\langle \phi_0\,,\,\phi_0\rangle} = \frac{\langle T_1\,,\,T_1\rangle}{\langle T_0\,,\,T_0\rangle} = \tfrac{1}{2}, \\
b_n &= \frac{\langle \phi_{n-1}\,,\,\phi_{n-1}\rangle}{\langle \phi_{n-2}\,,\,\phi_{n-2}\rangle} = \frac{\left\langle 2^{2-n}T_{n-1}\,,\,2^{2-n}T_{n-1}\right\rangle}{\left\langle 2^{3-n}T_{n-2}\,,\,2^{3-n}T_{n-2}\right\rangle} = \tfrac{1}{4} \quad (n > 2).
\end{aligned}$$

So

$$\begin{aligned}
\phi_0 &= 1, \\
\phi_1 &= x, \\
\phi_2 &= x\phi_1 - \tfrac{1}{2}\phi_0, \\
\phi_n &= x\phi_{n-1} - \tfrac{1}{4}\phi_{n-2}\ (n > 2).
\end{aligned}$$

Hence the recurrence (1.3) for T_n.

We may similarly derive the recurrences (1.6) for U_n and (1.12) for V_n and W_n, by using their respective weight functions to obtain the appropriate a_n and b_n (see Problem 5).

4.5 Rodrigues' formulae and differential equations

If $\{\phi_i\}$ is a set of polynomials orthogonal on $[-1, 1]$ with respect to $w(x)$, with ϕ_i of degree i, then

$$\int_{-1}^{1} w(x)\phi_n(x)q_{n-1}(x)\,\mathrm{d}x = 0 \tag{4.22}$$

for any polynomial q_{n-1} of degree $n-1$.

Now suppose that $r_n(x)$ is an nth integral of $w(x)\phi_n(x)$, so that

$$r_n^{(n)}(x) = w(x)\phi_n(x). \tag{4.23}$$

Then (4.22) gives, on integration by parts,

$$\begin{aligned}
0 &= \int_{-1}^{1} r_n^{(n)}(x)q_{n-1}(x)\,\mathrm{d}x \\
&= \left[r^{(n-1)}(x)q_{n-1}(x)\right]_{-1}^{1} - \int_{-1}^{1} r_n^{(n-1)}(x)q'_{n-1}(x)\,\mathrm{d}x \\
&= \left[r^{(n-1)}(x)q_{n-1}(x) - r^{(n-2)}(x)q'_{n-1}(x)\right]_{-1}^{1} + \int_{-1}^{1} r_n^{(n-2)}(x)q''_{n-1}(x)\,\mathrm{d}x \\
&= \cdots \\
&= \left[r^{(n-1)}(x)q_{n-1}(x) - r^{(n-2)}(x)q'_{n-1}(x) + \cdots + (-1)^{n-1}r_n(x)q_{n-1}^{(n-1)}(x)\right]_{-1}^{1} + \\
&\qquad + (-1)^n \int_{-1}^{1} r_n(x)q_{n-1}^{(n)}(x)\,\mathrm{d}x.
\end{aligned}$$

Hence, since $q_{n-1}^{(n)}(x) \equiv 0$, it follows that

$$\left[r^{(n-1)}(x)q_{n-1}(x) - r^{(n-2)}(x)q'_{n-1}(x) + \cdots + (-1)^{n-1}r_n(x)q_{n-1}^{(n-1)}(x)\right]_{-1}^{1} = 0 \tag{4.24}$$

for any polynomial q_{n-1} of degree $n-1$.

Now $\phi_n^{(n+1)}(x) \equiv 0$, since ϕ_n is of degree n; hence, because of (4.23), r_n is a solution of the $(2n+1)$st order homogeneous differential equation

$$\frac{\mathrm{d}^{n+1}}{\mathrm{d}x^{n+1}}\left(w(x)^{-1}\frac{\mathrm{d}^n}{\mathrm{d}x^n}r_n(x)\right) = 0. \tag{4.25}$$

An arbitrary polynomial of degree $n-1$ may be added to r_n, without affecting the truth of (4.23) and (4.25). Hence we may without loss of generality arrange that $r_n(-1) = r'_n(-1) = \cdots = r_n^{(n-1)}(-1) = 0$, when the fact that (4.24) is valid for all q_{n-1} implies that $r_n(+1) = r'_n(+1) = \cdots = r_n^{(n-1)}(+1) = 0$, so that r_n satisfies the $2n$ homogeneous boundary conditions

$$r_n(\pm 1) = r'_n(\pm 1) = \cdots = r_n^{(n-1)}(\pm 1) = 0. \tag{4.26}$$

One function satisfying (4.26), for any real $\alpha > -1$, is

$$r_n(x) = (1 - x^2)^{n+\alpha}. \tag{4.27}$$

If we then choose

$$w(x) = (1 - x^2)^{\alpha} \tag{4.28}$$

then (4.25) is satisfied, and $r_n^{(n)}$ is of the form (4.23) with $\phi_n(x)$ a polynomial of degree n.

Since ϕ_n, as defined by (4.22), is unique apart from a multiplicative constant, it follows from (4.25), (4.27) and (4.28) that (for $\alpha > -1$)

$$\phi_n(x) = P_n^{(\alpha)}(x) := c_n \frac{1}{(1 - x^2)^{\alpha}} \frac{\mathrm{d}^n}{\mathrm{d}x^n}(1 - x^2)^{n+\alpha}, \tag{4.29}$$

where c_n is a constant, defines a system of polynomials $\{P_n^{(\alpha)}(x)\}$ orthogonal with respect to $w(x) = (1 - x^2)^{\alpha}$ on $[-1, 1]$. These polynomials are known as the *ultraspherical* (or *Gegenbauer*) *polynomials*, and the formula (4.29) for them is known as *Rodrigues' formula*.

It immediately follows that the Chebyshev polynomials of the first and second kinds are ultraspherical polynomials and, by comparing their leading coefficients with those in (4.29), we may readily deduce (see Problem 12) that, taking $\alpha = -\frac{1}{2}$ and $\alpha = +\frac{1}{2}$,

$$T_n(x) = \frac{(-1)^n\, 2^n\, n!}{(2n)!}(1 - x^2)^{\frac{1}{2}} \frac{\mathrm{d}^n}{\mathrm{d}x^n}(1 - x^2)^{n-\frac{1}{2}}, \tag{4.30}$$

$$U_n(x) = \frac{(-1)^n\, 2^n\, (n+1)!}{(2n+1)!}(1 - x^2)^{-\frac{1}{2}} \frac{\mathrm{d}^n}{\mathrm{d}x^n}(1 - x^2)^{n+\frac{1}{2}}. \tag{4.31}$$

(In the standard notation for Gegenbauer polynomials, as in Abramowitz and Stegun's *Handbook of Mathematical Functions* (1964) for example, $P_n^{(\alpha)}(x)$ is written as $C_n^{\alpha+\frac{1}{2}}(x)$, so that $T_n(x)$ is proportional to $C_n^{(0)}(x)$ and $U_n(x)$ to $C_n^{(1)}(x)$.)

The well-known Legendre polynomials $P_n(x)$, which are orthogonal with weight unity, are ultraspherical polynomials for $\alpha = 0$ and are given by

$$P_n(x) = \frac{(-1)^n 2^{-n}}{n!} \frac{\mathrm{d}^n}{\mathrm{d}x^n}(1 - x^2)^n. \tag{4.32}$$

Note that the Chebyshev polynomials of third and fourth kinds are not ultraspherical polynomials, but only Jacobi polynomials. Their Rodrigues' formulae are

$$V_n(x) = \frac{(-1)^n\, 2^n\, n!}{(2n)!} \left(\frac{1-x}{1+x}\right)^{\frac{1}{2}} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \left\{(1-x^2)^n \left(\frac{1+x}{1-x}\right)^{\frac{1}{2}}\right\}, \tag{4.33a}$$

$$W_n(x) = \frac{(-1)^n\, 2^n\, n!}{(2n)!} \left(\frac{1+x}{1-x}\right)^{\frac{1}{2}} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \left\{(1-x^2)^n \left(\frac{1-x}{1+x}\right)^{\frac{1}{2}}\right\}. \tag{4.33b}$$

From the general formula (4.29) it can be verified by substitution (see Problem 13) that $P_n^{(\alpha)}(x)$ is a solution of the second-order differential equation

$$(1-x^2)y'' - 2(\alpha+1)xy' + n(n+2\alpha+1)y = 0. \tag{4.34}$$

Thus $T_n(x)$, $U_n(x)$, $P_n(x)$ are solutions of

$$(1-x^2)y'' - xy' + n^2y = 0 \quad (\alpha = -\tfrac{1}{2}), \tag{4.35a}$$

$$(1-x^2)y'' - 3xy' + n(n+2)y = 0 \quad (\alpha = \tfrac{1}{2}), \tag{4.35b}$$

$$(1-x^2)y'' - 2xy' + n(n+1)y = 0 \quad (\alpha = 0), \tag{4.35c}$$

respectively.

The differential equations satisfied by $V_n(x)$ and $W_n(x)$ are, respectively,

$$(1-x^2)y'' - (2x-1)y' + n(n+1)y = 0, \tag{4.36a}$$

$$(1-x^2)y'' - (2x+1)y' + n(n+1)y = 0. \tag{4.36b}$$

4.6 Discrete orthogonality of Chebyshev polynomials

It is always possible to convert a (continuous) orthogonality relationship, as defined in Definition 4.1, into a discrete orthogonality relationship simply by replacing the integral with a summation. In general, of course, the result is only approximately true. However, where trigonometric functions or Chebyshev polynomials are involved, there are many cases in which the discrete orthogonality can be shown to hold exactly. We give here a few of the discrete orthogonality relations that exist between the four kinds of Chebyshev polynomials. Further relations are given by Mason & Venturino (1996) (see also Problem 14).

4.6.1 First-kind polynomials

Consider the sum

$$s_n^{(1)}(\theta) = \sum_{k=1}^{n+1} \cos(k-\tfrac{1}{2})\theta = \cos\tfrac{1}{2}\theta + \cos\tfrac{3}{2}\theta + \cdots + \cos(n+\tfrac{1}{2})\theta. \tag{4.37}$$

By summing the arithmetic series

$$z^{\frac{1}{2}}(1 + z + z^2 + \cdots + z^n),$$

substituting $z = e^{i\theta}$ and taking the real part of the result, it is easily verified (see Problem 4) that

$$s_n^{(1)}(\theta) = \frac{\sin(n+1)\theta}{2\sin\frac{1}{2}\theta} \tag{4.38}$$

and hence that $s_n^{(1)}(\theta)$ vanishes when $\theta = \dfrac{r\pi}{n+1}$, for integers r in the range $0 < r < 2(n+1)$. Further, we can see directly from (4.37) that

$$s_n^{(1)}(0) = n+1, \quad s_n^{(1)}(2\pi) = -(n+1). \tag{4.39}$$

Now consider

$$a_{ij} = \sum_{k=1}^{n+1} T_i(x_k)T_j(x_k) \quad (0 \le i, j \le n) \tag{4.40}$$

where x_k are the zeros of $T_{n+1}(x)$, namely

$$x_k = \cos\theta_k, \quad \theta_k = \frac{(k-\frac{1}{2})\pi}{n+1}. \tag{4.41}$$

Then

$$\begin{aligned} a_{ij} &= \sum_{k=1}^{n+1} \cos i\theta_k \cos j\theta_k \\ &= \tfrac{1}{2}\sum_{k=1}^{n+1} [\cos(i+j)\theta_k + \cos(i-j)\theta_k] \\ &= \tfrac{1}{2}\left[s_n^{(1)}\left(\frac{(i+j)\pi}{n+1}\right) + s_n^{(1)}\left(\frac{(i-j)\pi}{n+1}\right)\right]. \end{aligned}$$

Hence

$$a_{ij} = 0 \quad (i \ne j;\ i, j \le n), \tag{4.42a}$$

while, using (4.39),

$$a_{ii} = \tfrac{1}{2}(n+1) \quad (i \ne 0;\ i \le n) \tag{4.42b}$$

and

$$a_{00} = n+1. \tag{4.42c}$$

It follows from (4.42a) that the polynomials $\{T_i(x),\ i = 0, 1, \ldots, n\}$ are orthogonal over the discrete point set $\{x_k\}$ consisting of the zeros of $T_{n+1}(x)$. Specifically, the orthogonality is defined for the discrete inner product

$$\langle u\,,\ v\rangle = \sum_{k=1}^{n+1} u(x_k)v(x_k) \tag{4.43}$$

in the form

$$\langle T_i\,,\,T_j\rangle = 0 \quad (i \neq j;\; i,j \leq n),$$

with $\langle T_0\,,\,T_0\rangle = n+1$ and $\langle T_i\,,\,T_i\rangle = \frac{1}{2}(n+1)$ $(0 < i \leq n)$.

This is not the only discrete orthogonality property of $\{T_i\}$. Indeed, by considering instead of (4.37) the sum

$$s_n^{(2)}(\theta) = \sum_{k=0}^{n}{}'' \cos k\theta = \tfrac{1}{2}\sin n\theta\,\cot\tfrac{1}{2}\theta \quad (n > 0)$$

(see Problem 4), where the double dash in $\sum''$ denotes that both first and last terms in the sum are to be halved, we deduce that

$$s_n^{(2)}(r\pi/n) = 0$$

for $0 < r < 2n$, while

$$s_n^{(2)}(0) = s_n^{(2)}(2\pi) = n.$$

If we now consider the extrema y_k of $T_n(x)$, namely

$$y_k = \cos\phi_k, \quad \phi_k = \frac{k\pi}{n} \;(k = 0, 1, \ldots, n) \tag{4.44}$$

(note that these $\{y_k\}$ are also the zeros of $U_{n-1}(x)$ together with the end points ± 1), and define

$$b_{ij} = \sum_{k=0}^{n}{}'' T_i(y_k)T_j(y_k), \tag{4.45}$$

then we have

$$b_{ij} = 0 \quad (i \neq j;\; i,j \leq n) \tag{4.46a}$$

$$b_{ii} = \tfrac{1}{2}n \quad (0 < i < n) \tag{4.46b}$$

$$b_{00} = b_{nn} = n. \tag{4.46c}$$

In this case the inner product is

$$\langle u\,,\,v\rangle = \sum_{k=0}^{n}{}'' u(y_k)v(y_k) \tag{4.47}$$

and we have again

$$\langle T_i\,,\,T_j\rangle = 0 \quad (i \neq j;\; i,j \leq n),$$

but this time with $\langle T_0\,,\,T_0\rangle = \langle T_n\,,\,T_n\rangle = n$ and $\langle T_i\,,\,T_i\rangle = \frac{1}{2}n$, $(0 < i < n)$.

4.6.2 Second-kind polynomials

In a similar way, we can establish a pair of discrete orthogonality relationships for the weighted second-kind polynomials $\{\sqrt{1-x^2}\,U_i(x)\}$ corresponding to the point sets $\{x_k\}$ and $\{y_k\}$ defined in (4.41) and (4.44)).

Define

$$a_{ij}^{(2)} = \sum_{k=1}^{n+1}(1-x_k^2)U_i(x_k)U_j(x_k) \quad (0 \le i,j \le n) \tag{4.48}$$

where $\{x_k\}$ are zeros of $T_{n+1}(x)$. Then we note that

$$\begin{aligned} a_{ij}^{(2)} &= \sum_{k=1}^{n+1} \sin(i+1)\theta_k \, \sin(j+1)\theta_k \\ &= \tfrac{1}{2}\sum_{k=1}^{n+1}[\cos(i-j)\theta_k - \cos(i+j+2)\theta_k] \\ &= \tfrac{1}{2}\left[s_n^{(1)}\left(\frac{(i-j)\pi}{n+1}\right) - s_n^{(1)}\left(\frac{(i+j+2)\pi}{n+1}\right)\right]. \end{aligned}$$

Hence

$$a_{ij}^{(2)} = 0 \quad (i \ne j;\ 0 \le i,j \le n) \tag{4.49a}$$

and

$$a_{ii}^{(2)} = \tfrac{1}{2}(n+1) \quad (0 \le i < n), \tag{4.49b}$$

while

$$a_{nn}^{(2)} = n+1. \tag{4.49c}$$

Thus $\{\sqrt{1-x^2}\,U_i(x),\ i = 0,1,\ldots,n\}$ are orthogonal for the inner product (4.43).

Similarly, considering the zeros $\{y_k\}$ of $(1-x^2)U_{n-1}(x)$,

$$\begin{aligned} b_{ij}^{(2)} &= \sum_{k=0}^{n}{}'' (1-y_k^2)U_i(y_k)U_j(y_k) \\ &= \sum_{k=0}^{n}{}'' \sin(i+1)\phi_k \, \sin(j+1)\phi_k. \end{aligned} \tag{4.50}$$

Then

$$b_{ij}^{(2)} = 0 \quad (i \ne j;\ i,j \le n-1) \tag{4.51a}$$

$$b_{ii}^{(2)} = \tfrac{1}{2}n \quad (0 \le i < n-1) \tag{4.51b}$$

$$b_{n-1,n-1}^{(2)} = 0 \tag{4.51c}$$

and $\{\sqrt{1-x^2}\,U_i(x),\ i = 0,1,\ldots,n-1\}$ are orthogonal for the inner product (4.47).

4.6.3 Third- and fourth-kind polynomials

Surprisingly, perhaps, the same discrete abscissae and inner products (4.43) and (4.47) provide orthogonality for the weighted third- and fourth-kind polynomials

$$\{\sqrt{1+x}\,V_i(x)\},\ \{\sqrt{1-x}\,W_i(x)\}.$$

For we have

$$\begin{aligned} a_{ij}^{(3)} &= \sum_{k=1}^{n+1}(1+x_k)V_i(x_k)V_j(x_k) \quad (0 \le i,j \le n) \\ &= 2\sum_{k=1}^{n+1}\cos(i+\tfrac{1}{2})\theta_k\ \cos(j+\tfrac{1}{2})\theta_k \\ &= \sum_{k=1}^{n+1}[\cos(i+j+1)\theta_k + \cos(i-j)\theta_k], \end{aligned}$$

giving us

$$a_{ij}^{(3)} = 0 \quad (i \ne j;\ i,j \le n) \tag{4.52a}$$

$$a_{ii}^{(3)} = n+1 \quad (0 \le i \le n), \tag{4.52b}$$

while

$$\begin{aligned} b_{ij}^{(3)} &= \sideset{}{''}\sum_{k=0}^{n}(1+y_k)V_i(y_k)V_j(y_k) \quad (0 \le i,j \le n) \\ &= 2\ \sideset{}{''}\sum_{k=0}^{n}\cos(i+\tfrac{1}{2})\phi_k\ \cos(j+\tfrac{1}{2})\phi_k \\ &= \sideset{}{''}\sum_{k=0}^{n}[\cos(i+j+1)\phi_k + \cos(i-j)\phi_k], \end{aligned}$$

giving

$$b_{ij}^{(3)} = 0 \quad (i \ne j;\ i,j \le n) \tag{4.53a}$$

$$b_{ii}^{(3)} = n \quad (0 \le i \le n). \tag{4.53b}$$

The same formulae (4.52)–(4.53) hold for $a_{ij}^{(4)}$ and $b_{ij}^{(4)}$, where

$$a_{ij}^{(4)} = \sum_{k=1}^{n+1}(1-x_k)W_i(x_k)W_j(x_k) \text{ and } b_{ij}^{(4)} = \sideset{}{''}\sum_{k=0}^{n}(1-y_k)W_i(y_k)W_j(y_k). \tag{4.54}$$

4.7 Discrete Chebyshev transforms and the fast Fourier transform

Using the values of a function $f(x)$ at the extrema $\{y_k\}$ of $T_n(x)$, which are also the zeros of $(1 - x^2)U_{n-1}(x)$, given as in (4.44) by

$$y_k = \cos\frac{k\pi}{n} \quad (k = 0, \ldots, n), \tag{4.55}$$

we can define a *discrete Chebyshev transform* $\hat{f}(x)$, defined at these same points only, by the formula

$$\hat{f}(y_k) := \sqrt{\tfrac{2}{n}} \sideset{}{''}\sum_{j=0}^{n} T_k(y_j)f(y_j) \quad (k = 0, \ldots, n). \tag{4.56}$$

These values $\hat{f}(y_k)$ are in fact proportional to the coefficients in the interpolant of $f(y_k)$ by a sum of Chebyshev polynomials — see Section 6.3.2.

Using the discrete orthogonality relation (4.45), namely

$$\sideset{}{''}\sum_{k=0}^{n} T_i(y_k)T_j(y_k) = \begin{cases} 0 & (i \neq j;\ i, j \leq n), \\ \frac{1}{2}n & (0 < i = j < n), \\ n & (i = j = 0 \text{ or } n), \end{cases} \tag{4.57}$$

we can easily deduce that the inverse transform is given by

$$f(y_j) = \sqrt{\tfrac{2}{n}} \sideset{}{''}\sum_{k=0}^{n} T_k(y_j)\hat{f}(y_k) \quad (j = 0, \ldots, n). \tag{4.58}$$

In fact, since

$$T_k(y_j) = \cos\frac{jk\pi}{n} = T_j(y_k),$$

which is symmetric in j and k, it is clear that the discrete Chebyshev transform is self-inverse.

It is possible to define other forms of discrete Chebyshev transform, based on any of the other discrete orthogonality relations detailed in Section 4.6.

The discrete Chebyshev transform defined here is intimately connected with the discrete Fourier (cosine) transform. Defining

$$\phi_k = \frac{k\pi}{n}$$

(the zeros of $\sin n\theta$) and

$$g(\theta) = f(\cos\theta), \quad \hat{g}(\theta) = \hat{f}(\cos\theta),$$

the formula (4.56) converts to the form

$$\hat{g}(\phi_k) := \sqrt{\tfrac{2}{n}} \sum_{j=0}^{n}{}'' \cos\frac{jk\pi}{n} g(\phi_j) \quad (k = 0, \ldots, n). \tag{4.59}$$

Since $\cos\theta$ and therefore $g(\theta)$ are even and 2π-periodic functions of θ, (4.59) has alternative equivalent expressions

$$\hat{g}\left(\frac{k\pi}{n}\right) = \sqrt{\tfrac{1}{2n}} \sum_{j=-n}^{n}{}'' \cos\frac{jk\pi}{n}\, g\left(\frac{j\pi}{n}\right) \quad (k = -n, \ldots, n) \tag{4.60}$$

or

$$\hat{g}\left(\frac{k\pi}{n}\right) = \sqrt{\tfrac{1}{2n}} \sum_{j=-n}^{n}{}'' \exp\frac{\mathrm{i}jk\pi}{n}\, g\left(\frac{j\pi}{n}\right) \quad (k = -n, \ldots, n) \tag{4.61}$$

or

$$\hat{g}\left(\frac{k\pi}{n}\right) = \sqrt{\tfrac{1}{2n}} \sum_{j=0}^{2n-1} \exp\frac{\mathrm{i}jk\pi}{n}\, g\left(\frac{j\pi}{n}\right) \quad (k = 0, \ldots, 2n-1). \tag{4.62}$$

The last formulae (4.61) and (4.62) in fact define the general discrete Fourier transform, applicable to functions $g(\theta)$ that are periodic but not necessarily even, whose inverse is the complex conjugate transform

$$g\left(\frac{j\pi}{n}\right) = \sqrt{\tfrac{1}{2n}} \sum_{k=-n}^{n}{}'' \exp\frac{-\mathrm{i}jk\pi}{n}\, \hat{g}\left(\frac{k\pi}{n}\right) \quad (j = -n, \ldots, n) \tag{4.63}$$

or

$$g\left(\frac{j\pi}{n}\right) = \sqrt{\tfrac{1}{2n}} \sum_{k=0}^{2n-1} \exp\frac{-\mathrm{i}jk\pi}{n}\, \hat{g}\left(\frac{k\pi}{n}\right) \quad (j = 0, \ldots, 2n-1). \tag{4.64}$$

4.7.1 The fast Fourier transform

Evaluation of (4.56) or (4.58) for a particular value of k or j, respectively, requires a number $O(n)$ of arithmetic operations; the algorithm described in Section 2.4.1 is probably the most efficient. If we require their values to be calculated for all values of k or j, however, use of this scheme would call for $O(n^2)$ operations in all, whereas it is possible to achieve the same results in $O(n \log n)$ operations (at the slight cost of working in complex arithmetic rather than real arithmetic, even though the final result is known to be real) by converting the Chebyshev transform to the equivalent Fourier transform (4.62) or (4.64), and then computing its $2n$ values simultaneously by means of the so-called *fast Fourier transform* (FFT) algorithm (Cooley & Tukey 1965, Gentleman & Sande 1966). The required $n+1$ values of the Chebyshev transform may then be extracted. (The remaining $n-1$ computed results will be redundant, by reason of the symmetry of g and $\hat{g}$.)

While there are versions of this algorithm that apply when n is a product of any small prime factors (Kolba & Parks 1977, Burrus & Eschenbacher 1981, for instance), it is easiest to describe it for the original and most useful case where n is a power of 2; say $n = 2^m$. Then, separating the even and odd terms of the summation, (4.62) becomes

$$\begin{aligned}\hat{g}\left(\frac{k\pi}{n}\right) &= \sqrt{\tfrac{1}{2n}} \sum_{\substack{j=0 \\ j \text{ even}}}^{2n-2} \exp\frac{\mathrm{i}jk\pi}{n}\, g\left(\frac{j\pi}{n}\right) + \sqrt{\tfrac{1}{2n}} \sum_{\substack{j=1 \\ j \text{ odd}}}^{2n-1} \exp\frac{\mathrm{i}jk\pi}{n}\, g\left(\frac{j\pi}{n}\right) \\ &= \sqrt{\tfrac{1}{2n}} \sum_{j=0}^{n-1} \exp\frac{2\mathrm{i}jk\pi}{n}\, g\left(\frac{2j\pi}{n}\right) + \\ &\qquad + \sqrt{\tfrac{1}{2n}} \exp\frac{\mathrm{i}k\pi}{n} \sum_{j=0}^{n-1} \exp\frac{2\mathrm{i}jk\pi}{n}\, g\left(\frac{(2j+1)\pi}{n}\right). \qquad (4.65a)\end{aligned}$$

while

$$\begin{aligned}\hat{g}\left(\frac{(k+n)\pi}{n}\right) &= \sqrt{\tfrac{1}{2n}} \sum_{j=0}^{n-1} \exp\frac{2\mathrm{i}jk\pi}{n}\, g\left(\frac{2j\pi}{n}\right) - \\ &\qquad - \sqrt{\tfrac{1}{2n}} \exp\frac{\mathrm{i}k\pi}{n} \sum_{j=0}^{n-1} \exp\frac{2\mathrm{i}jk\pi}{n}\, g\left(\frac{(2j+1)\pi}{n}\right). \qquad (4.65b)\end{aligned}$$

Now, if for $j = 0, \ldots, n-1$ we define

$$g_1\left(\frac{2j\pi}{n}\right) := g\left(\frac{2j\pi}{n}\right) \text{ and } g_2\left(\frac{2j\pi}{n}\right) := g\left(\frac{(2j+1)\pi}{n}\right),$$

we can further rewrite (4.65a) and (4.65b) as

$$\hat{g}\left(\frac{k\pi}{n}\right) = \tfrac{1}{\sqrt{2}}\, \hat{g}_1\left(\frac{2k\pi}{n}\right) + \tfrac{1}{\sqrt{2}} \exp\frac{\mathrm{i}k\pi}{n}\, \hat{g}_2\left(\frac{2k\pi}{n}\right), \qquad (4.66a)$$

$$\hat{g}\left(\frac{(k+n)\pi}{n}\right) = \tfrac{1}{\sqrt{2}}\, \hat{g}_1\left(\frac{2k\pi}{n}\right) - \tfrac{1}{\sqrt{2}} \exp\frac{\mathrm{i}k\pi}{n}\, \hat{g}_2\left(\frac{2k\pi}{n}\right), \qquad (4.66b)$$

$$(k = 0, \ldots, n-1)$$

where the discrete Fourier transforms from g_1 to $\hat{g}_1$ and from g_2 to $\hat{g}_2$ each take a set of n values into another set of n values, whereas that from g to $\hat{g}$ takes $2n$ values into $2n$.

Thus, once we have performed the two transforms of order n, it requires fewer than Kn further arithmetic operations (where K denotes a small fixed integer) to compute the transform of order $2n$. Similarly it requires fewer than $2 \times K\frac{n}{2} = Kn$ operations to derive these two transforms of order n from four transforms of order $n/2$, fewer than $4 \times K\frac{n}{4} = Kn$ operations to derive these four transforms of order $n/2$ from eight transforms of order $n/4$, and so on. If $n = 2^m$, therefore, (a transform of order 1 being just the identity and therefore trivial) the discrete Fourier transform of order $2n$ may be performed in $(m+1)$ stages, each requiring fewer than Kn operations, so that the total number of operations is less than $(m+1)Kn = O(n \log n)$, as claimed above.

We do not propose to discuss in detail how this computation is best organised, but refer the reader to the extensive published literature (Canuto et al. 1988, van Loan 1992, for instance). Reliable off-the-peg implementations of a fast Fourier transform algorithm can be found in any comprehensive numerical subroutine library.

4.8 Discrete data fitting by orthogonal polynomials: the Forsythe–Clenshaw method

In this section we consider a least-squares/orthogonal polynomial method, in which Chebyshev polynomials fulfil what is essentially a supporting role. However, this is one of the most versatile polynomial approximation algorithms available, and the use of Chebyshev polynomials makes the resulting approximations much easier to use and compute. Moreover, the algorithm, in its Chebyshev polynomial form, is an essential tool in the solution of multivariate data-fitting problems for data on families of lines or curves.

We saw in Section 4.6 that an inner product may be defined on a discrete data set just as well as on a continuum, and in (4.43) we defined such an inner product based on Chebyshev polynomial zeros. However, we are frequently given a set of arbitrarily spaced data abscissae

$$x = x_k \quad (k = 1, \ldots, m), \tag{4.67}$$

and asked to approximate in a least-squares sense the corresponding ordinates

$$y = y_k$$

by a polynomial of degree n, where the number $(n+1)$ of free parameters is no more than the number m of given data — typically much smaller.

Now

$$\langle u\,,\,v\rangle = \sum_{k=1}^{m} w_k u(x_k) v(x_k) \tag{4.68}$$

defines an inner product over the points (4.67), where $\{w_k\}$ is a specified set of positive weights to be applied to the data. From Corollary 4.1B, the best polynomial approximation of degree n in the least-squares sense on the point set (4.67) is therefore

$$p_n^B = \sum_{i=0}^{n} c_i \phi_i(x) \tag{4.69}$$

where $\{\phi_i\}$ are orthogonal polynomials defined by the recurrence (4.21) with the inner product (4.68) and where

$$\begin{aligned} c_i &= \frac{\langle y\,,\,\phi_i\rangle}{\langle \phi_i\,,\,\phi_i\rangle} \\ &= \frac{\displaystyle\sum_{k=1}^{m} w_k y_k \phi_i(x_k)}{\displaystyle\sum_{k=1}^{m} w_k [\phi_i(x_k)]^2}. \end{aligned} \tag{4.70}$$

This is precisely the algorithm proposed by Forsythe (1957) for approximating discrete data y_k at arbitrary points x_k $(k = 1, \ldots, n)$.

The Forsythe algorithm, as we have described it so far, does not explicitly involve Chebyshev polynomials (or, for that matter, any other well-known set of orthogonal polynomials). However, if the data are distributed uniformly and very densely over an interval of x, $[-1, 1]$ say, then we expect the resulting polynomials to be very similar to conventional orthogonal polynomials defined on the continuum $[-1, 1]$. For example, if all the w_k are equal to unity, then $\{\phi_k\}$ should closely resemble the Legendre polynomials (orthogonal with respect to $w(x) = 1$), and if

$$w_k = (1 - x_k^2)^{-\frac{1}{2}}$$

then $\{\phi_k\}$ should resemble the Chebyshev polynomials of the first kind. In spite of this resemblance, we cannot simply use the Legendre or Chebyshev polynomials in place of the polynomials ϕ_k in (4.69) and (4.70), since on these points they are only approximately orthogonal, not exactly, and so we have to consider some other approach.

The goal is a formula for p_n^B based on Chebyshev polynomials of the first kind, say, in the form

$$p_n^B(x) = {\sum_{i=0}^{n}}' d_i^{(n)} T_i(x), \tag{4.71}$$

the new set of coefficients $d_i^{(n)}$ being chosen so that (4.71) is identical to (4.69). This form has the advantage over (4.69) that the basis $\{T_i(x)\}$ is independent of the abscissae (4.67) and therefore more convenient for repeated computation of $p_n^B(x)$. (This is a very useful step in the development of multivariate polynomial approximations on lines or curves of data.) An efficient algorithm for deriving the coefficients $d_i^{(n)}$ from the coefficients c_i is due to Clenshaw (1959/1960); it makes use of the recurrence relations for Chebyshev polynomials as well as those for the discretely orthogonal polynomials $\{\phi_i\}$. We shall not give more details here.

4.8.1 Bivariate discrete data fitting on or near a family of lines or curves

Formula (4.71), which gives an approximation p_n^B to data unequally spaced along one line, may readily be extended to much more general situations in two or more dimensions (Clenshaw & Hayes 1965). In two dimensions, for example, suppose that data are given at unequally-spaced and different locations on each of a family of lines parallel to the x-axis, say at the points

$$(x_{k\ell}, y_\ell), \quad k = 1, \ldots, m_{1\ell};\ \ell = 1, \ldots, m_2.$$

We suppose too that all of these points lie within the square $[-1, 1] \times [-1, 1]$. Then the data on each line $y = y_\ell$ may be approximated, using Clenshaw's algorithm, in the form (4.71), giving us a set of approximations[1]

$$p_{n_1,\ell}^B(x) = \sum_{i=0}^{n_1}{}' d_{i\ell}^{(n_1)} T_i(x), \quad \ell = 1, \ldots, m_2. \tag{4.72}$$

The set of ith coefficients $d_{i\ell}^{(n_1)}$, $\ell = 1, \ldots, m_2$, may then be treated as data on a line parallel to the y-axis and may be approximated in a similar manner, for each i from 0 to n_1, giving approximations

$$d_i(y) = \sum_{j=0}^{n_2}{}' d_{ij}^{(n_1,n_2)} T_j(y). \tag{4.73}$$

We thus arrive at the overall approximation

$$\sum_{i=0}^{n_1}{}' \sum_{j=0}^{n_2}{}' d_{ij}^{(n_1,n_2)} T_i(x) T_j(y). \tag{4.74}$$

If m_1, m_2, n_1, n_2 all $= O(n)$, then the algorithm involves $O(n^4)$ operations—compared with $O(n^3)$ for a meshed data polynomial (tensor product) algorithm. It is important (Clenshaw & Hayes 1965) to ensure that there are data located close to $x = \pm 1$ and $y = \pm 1$, if necessary by changing variables to transform boundary curves into straight lines.

[1]If the number of data points on any line is less than the number of degrees of freedom, $m_{1\ell} \leq n_1$, then instead of approximating we can interpolate with a polynomial of minimal degree by requiring that $d_{i\ell}^{(n_1)} = 0$ for $i \geq m_{1\ell}$.

This algorithm has been extended further by Bennell & Mason (1991) to data on a family of curves. The procedure is to run secant lines across the family of curved lines and interpolate the data on each curve to give values at the intersections, which are then approximated by using the method just described. The algorithm involves only about twice as many operations as that for data on lines, which would appear very satisfactory.

Other writers have developed similar algorithms. For example, Anderson et al. (1995) fit data lying 'near' a family of lines, using an iteration based on estimating values on the lines from the neighbouring data.

Algorithms such as these have considerable potential in higher dimensions. An application to modelling the surfaces of human teeth has been successfully carried out by Jovanovski (1999).

4.9 Orthogonality in the complex plane

Formulae for Chebyshev polynomials in terms of a complex variable z have been given in Section 1.4; we repeat them here for convenience.

Given any complex number z, we define the related complex number w to be such that

$$z = \tfrac{1}{2}(w + w^{-1}). \tag{4.75}$$

Unless z lies on the real interval $[-1, 1]$, this equation for w has two solutions

$$w = z \pm \sqrt{z^2 - 1}, \tag{4.76}$$

one of which has $|w| > 1$ and one has $|w| < 1$; we choose the one with $|w| > 1$. Then we have:

$$T_n(z) = \tfrac{1}{2}(w^n + w^{-n}); \tag{4.77}$$

$$U_n(z) = \frac{w^{n+1} - w^{-n-1}}{w - w^{-1}}; \tag{4.78}$$

$$V_n(z) = \frac{w^{n+\frac{1}{2}} + w^{-n-\frac{1}{2}}}{w^{\frac{1}{2}} + w^{-\frac{1}{2}}} = \frac{w^{n+1} + w^{-n}}{w + 1}; \tag{4.79}$$

$$W_n(z) = \frac{w^{n+\frac{1}{2}} - w^{-n-\frac{1}{2}}}{w^{\frac{1}{2}} - w^{-\frac{1}{2}}} = \frac{w^{n+1} - w^{-n}}{w - 1}. \tag{4.80}$$

For any $r > 1$, the elliptical contour E_r given by

$$E_r := \left\{ z : \left| z + \sqrt{z^2 - 1} \right| = r \right\} \quad (r > 1) \tag{4.81}$$

has foci at $z = \pm 1$, and is the image under (4.75) of the circle

$$C_r := \{ w : |w| = r \} \tag{4.82}$$

of radius r.

The Chebyshev polynomials have useful orthogonality properties on the ellipse E_r. In order to describe them, we need first to extend our definition (Definition 4.2) of the idea of an inner product to allow for functions that take complex values. Let $\overline{z}$ denote the complex conjugate of z.

Definition 4.3 *An* inner product $\langle \cdot \,,\, \cdot \rangle$ *is defined as a bilinear function of elements $f, g, h, \ldots$ of a vector space that satisfies the axioms:*

1. $\langle f \,,\, f \rangle$ *is real and* ≥ 0*, with equality if and only if* $f \equiv 0$*;*
2. $\langle f \,,\, g \rangle = \overline{\langle g \,,\, f \rangle}$ *(note the complex conjugate);*
3. $\langle f + g \,,\, h \rangle = \langle f \,,\, h \rangle + \langle g \,,\, h \rangle$*;*
4. $\langle \alpha f \,,\, g \rangle = \alpha \langle f \,,\, g \rangle$ *for any scalar* α*.*
   (Hence $\langle f \,,\, \alpha g \rangle = \overline{\alpha} \langle f \,,\, g \rangle$*.)*

This definition agrees with the earlier one if everything is real.

Now define an inner product

$$\langle f \,,\, g \rangle := \oint_{E_r} f(z)\overline{g(z)}\, |\mu(z)|\, |\mathrm{d}z| \,, \tag{4.83}$$

where $\mu(z)$ is a weighting function ($|\mu(z)|$ is real and positive) and $\oint \cdots |\mathrm{d}z|$ denotes integration with respect to arc length around the ellipse in an anticlockwise direction. This inner product corresponds to a norm $\|\cdot\|_2$ on E_r defined by

$$\|f\|_2^2 := \langle f \,,\, f \rangle = \oint_{E_r} |f(z)|^2\, |\mu(z)|\, |\mathrm{d}z| \,. \tag{4.84}$$

Then we can show, using this inner product, that

$$\langle T_m \,,\, T_n \rangle = \begin{cases} 0 & (m \neq n) \\ 2\pi & (m = n = 0) \\ \frac{1}{2}\pi(r^{2n} + r^{-2n}) & (m = n > 0) \end{cases}$$
$$\text{if } \mu(z) = \frac{1}{\sqrt{1 - z^2}}, \tag{4.85a}$$

$$\langle U_m \,,\, U_n \rangle = \begin{cases} 0 & (m \neq n) \\ \frac{1}{2}\pi(r^{2n+2} + r^{-2n-2}) & (m = n \geq 0) \end{cases}$$
$$\text{if } \mu(z) = \sqrt{1 - z^2}, \tag{4.85b}$$

$$\langle V_m \,,\, V_n \rangle = \begin{cases} 0 & (m \neq n) \\ \pi(r^{2n+1} + r^{-2n-1}) & (m = n \geq 0) \end{cases}$$

$$\text{if } \mu(z) = \sqrt{\frac{1+z}{1-z}}, \tag{4.85c}$$

$$\langle W_m\,,\,W_n\rangle = \begin{cases} 0 & (m \neq n) \\ \pi(r^{2n+1} + r^{-2n-1}) & (m = n \geq 0) \end{cases}$$

$$\text{if } \mu(z) = \sqrt{\frac{1-z}{1+z}}. \tag{4.85d}$$

Proof: Taking the first of these orthogonalities, for example, we have

$$\begin{aligned}
z &= \tfrac{1}{2}(w + w^{-1}) = \tfrac{1}{2}(re^{i\theta} + r^{-1}e^{-i\theta}),\\
dz &= \tfrac{1}{2}(1 - w^{-2})\,dw = \tfrac{1}{2}i(re^{i\theta} - r^{-1}e^{-i\theta})\,d\theta,\\
T_m(z) &= \tfrac{1}{2}(w^m + w^{-m}) = \tfrac{1}{2}(r^m e^{im\theta} + r^{-m}e^{-im\theta}),\\
\mu(z) &= \frac{2}{\sqrt{2 - w^2 - w^{-2}}} = \frac{2}{\pm i(w - w^{-1})} = \pm\frac{2i}{re^{i\theta} - r^{-1}e^{-i\theta}},\\
|\mu(z)|\,|dz| &= d\theta,\\
T_m(z)\overline{T_n(z)} &= \tfrac{1}{4}(r^{m+n}e^{i(m-n)\theta} + r^{m-n}e^{i(m+n)\theta} +\\
&\quad + r^{-m+n}e^{-i(m+n)\theta} + r^{-m-n}e^{-i(m-n)\theta}).
\end{aligned}$$

Now

$$\int_0^{2\pi} e^{im\theta}\,d\theta = 0 \quad (m \neq 0), \quad = 2\pi \quad (m = 0).$$

Hence we easily show that $\langle T_m\,,\,T_n\rangle = 0$ for $m \neq n$ and is as stated for $m = n$. ••

The other results (4.85) may be proved similarly (Problem 19), after noting that

$$\begin{aligned}
\left|1 - z^2\right| &= \frac{1}{4\,|w|^2}\left|1 - w^2\right|^2,\\
|1 + z| &= \frac{1}{2\,|w|}\,|1 + w|^2,\\
|1 - z| &= \frac{1}{2\,|w|}\,|1 - w|^2.
\end{aligned}$$

Note that $\|f\|_2$, as defined in (4.84), as well as being a norm on the space of functions square integrable round the contour E_r, can be used as a norm on the space of functions which are continuous on this contour and analytic throughout its interior. This may be shown using the maximum-modulus theorem (Problem 21).

Alternatively, we may define an inner product (and a corresponding norm) over the whole of the interior D_r of the ellipse E_r

$$\langle f\,,\,g\rangle = \iint_{D_r} f(z)\overline{g(z)}\,|\mu(z)|\,dx\,dy, \tag{4.86}$$

where $z = x + \mathrm{i}y$. Remarkably, the Chebyshev polynomials are orthogonal with respect to this inner product too, defining $\mu(z)$ for each kind of polynomial as in equations (4.85).

Proof: Take the first-kind polynomials, for example, when $\mu(z) = 1/\sqrt{1-z^2}$. If $z = \frac{1}{2}(w + w^{-1})$ and $w = s\mathrm{e}^{\mathrm{i}\theta}$, then z runs over the whole ellipse D_r when s runs from 1 to r and θ runs from 0 to 2π. We have

$$\begin{aligned} x &= \tfrac{1}{2}(s + s^{-1})\cos\theta, \\ y &= \tfrac{1}{2}(s - s^{-1})\sin\theta, \end{aligned}$$

so that

$$\mathrm{d}x\ \mathrm{d}y = \frac{\partial(x,y)}{\partial(s,\theta)}\,\mathrm{d}s\ \mathrm{d}\theta$$

with

$$\begin{aligned} \frac{\partial(x,y)}{\partial(s,\theta)} &= \det\begin{pmatrix} \frac{1}{2}(1 - s^{-2})\cos\theta & -\frac{1}{2}(s + s^{-1})\sin\theta \\ +\frac{1}{2}(1 + s^{-2})\sin\theta & \frac{1}{2}(s - s^{-1})\cos\theta \end{pmatrix} \\ &= \frac{1}{4}\frac{(s^2 + 2s\cos\theta + 1)(s^2 - 2s\cos\theta + 1)}{s^3} \end{aligned}$$

while

$$|\mu(z)| = \frac{4s^2}{(s^2 + 2s\cos\theta + 1)(s^2 - 2s\cos\theta + 1)}$$

Thus,

$$\begin{aligned} \langle f\,,\,g\rangle &= \iint_{D_r} f(z)\overline{g(z)}\,|\mu(z)|\ \mathrm{d}x\ \mathrm{d}y \\ &= \int_{s=1}^{r}\int_{\theta=0}^{2\pi} f(z)\overline{g(z)}\,|\mu(z)|\,\frac{\partial(x,y)}{\partial(s,\theta)}\,\mathrm{d}s\ \mathrm{d}\theta \\ &= \int_{s=1}^{r}\int_{\theta=0}^{2\pi} f(z)\overline{g(z)}s^{-1}\,\mathrm{d}s\ \mathrm{d}\theta \\ &= \int_{s=1}^{r} s^{-1}\,\mathrm{d}s\left\{\int_{\theta=0}^{2\pi} f(z)\overline{g(z)}\,\mathrm{d}\theta\right\}. \end{aligned}$$

But the inner integral is simply the inner product around the ellipse E_s, which we have already shown to vanish if $f(z) = T_m(z)$ and $g(z) = T_n(z)$, with $m \neq n$. Therefore, the whole double integral vanishes, and $\langle T_m\,,\,T_n\rangle = 0$ for $m \neq n$. ●●

Orthogonality of the other three kinds of polynomial may be proved in the same way (Problem 20).

4.10 Problems for Chapter 4

1. Verify that the inner product (4.2) satisfies the axioms of Definition 4.2.

2. Using only the properties of an inner product listed in Definition 4.2, show that the norm defined by (4.4) satisfies the axioms of Definition 3.1.

3. If ϕ_n and ψ_n are two polynomials of degree n, each of which is orthogonal to every polynomial of degree less than n (over the same interval and with respect to the same weight function), show that $\phi_n(x)$ and $\psi_n(x)$ are proportional.

4. Derive the summations

$$\sum_{k=1}^{n+1} \cos(k - \tfrac{1}{2})\theta = \frac{\sin(n+1)\theta}{2\sin\frac{1}{2}\theta}$$

$$\sum_{k=1}^{n+1} \sin(k - \tfrac{1}{2})\theta = \frac{1 - \cos(n+1)\theta}{2\sin\frac{1}{2}\theta}$$

$$\sideset{}{''}\sum_{k=0}^{n} \cos k\theta = \tfrac{1}{2}\sin n\theta \cot\tfrac{1}{2}\theta$$

$$\sideset{}{''}\sum_{k=0}^{n} \sin k\theta = \tfrac{1}{2}(1 - \cos n\theta)\cot\tfrac{1}{2}\theta.$$

5. Using a similar analysis to that in Section 4.4, derive from (4.21) and the trigonometric formulae for $U_n(x)$, $V_n(x)$ and $W_n(x)$ the recurrence relations which are satisfied by these polynomials. Show that these relations coincide.

6. Using the recurrence (4.21), obtain formulae for the monic (Legendre) polynomials of degrees 0, 1, 2, 3, 4, which are orthogonal on $[-1, 1]$ with respect to $w(x) = 1$.

7. If $\{\phi_r\}$ is an orthogonal system on $[-1, 1]$, with ϕ_r a polynomial of exact degree r, prove that the zeros of ϕ_{r-1} separate those of ϕ_r; that is to say, between any two consecutive zeros of ϕ_r there lies a zero of ϕ_{r-1}. [Hint: Consider the signs of ϕ_r as $x \to +\infty$ and at the zeros of ϕ_{r-1}, using the recurrence (4.21).]

8. A simple alternative to the recurrence (4.21) for the generation of a system of monic orthogonal polynomials is the Gram–Schmidt orthogonalisation procedure:

   Given monic orthogonal polynomials $\phi_0, \phi_1, \ldots, \phi_{n-1}$, define ϕ_n in the form

$$\phi_n(x) = x^n + \sum_{k=0}^{n-1} c_k \phi_k(x)$$

and determine values of c_k such that ϕ_n is orthogonal to ϕ_0, ϕ_1, …, ϕ_{n-1}.

Use this recurrence to generate monic polynomials of degrees 0, 1, 2 orthogonal on $[-1, 1]$ with respect to $(1+x)^{-\frac{1}{2}}(1-x)^{\frac{1}{2}}$. What is the key disadvantage (in efficiency) of this algorithm, compared with the recurrence (4.21)?

9. By using the trigonometric formulae for $T_n(x)$ and $U_n(x)$, under the transformation $x = \cos\theta$, verify that these Chebyshev polynomials satisfy the respective differential equations (4.35a), (4.35b).

   Show similarly that $V_n(x)$ and $W_n(x)$ satisfy the differential equations (4.36a), (4.36b).

10. The second order differential equation (4.35a)

$$(1-x^2)y'' - xy' + n^2y = 0$$

   has $T_n(x)$ as one solution. Show that a second solution is $\sqrt{1-x^2}U_{n-1}(x)$. Find a second solution to (4.35b)

$$(1-x^2)y'' - 3xy' + n(n+2)y = 0,$$

   one solution of which is $U_n(x)$.

11. By substituting $T_n(x) = t_0 + t_1x + \cdots + t_nx^n$ into the differential equation that it satisfies, namely

$$(1-x^2)y'' - xy' + n^2y = 0,$$

   and equating coefficients of powers of x, show that $t_{n-1} = 0$ and

$$t_k(n^2 - k^2) + t_{k+2}(k+2)(k+1) = 0, \quad k = 0, \ldots, n-2.$$

   Deduce that

$$t_{n-2m} = (-1)^m \frac{n}{n-m}\binom{n-m}{m}2^{n-2m-1}$$

   where

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!}.$$

12. Writing

$$\frac{\mathrm{d}^r}{\mathrm{d}x^r}(1-x^2)^{n+\alpha} = (1-x^2)^{n-r+\alpha}(A_rx^r + \text{ lower degree terms}),$$

   show that $A_{r+1} = -(2n - r + 2\alpha)A_r$ and deduce that

$$A_n = (-1)^n(2n+2\alpha)(2n+2\alpha-1)\cdots(n+2\alpha).$$

   Hence, verify the formulae (4.30), (4.31) for $T_n(x)$, $U_n(x)$, determining the respective values of c_n in (4.29) by equating coefficients of x^n.

13. Verify that $P_n^{(\alpha)}(x)$, given by (4.29), is a solution of the second order equation

$$(1-x^2)y'' - 2(\alpha+1)xy' + n(n+2\alpha+1)y = 0.$$

[Hint: Write

$$\psi_n(x) := c_n^{-1}(1-x^2)^\alpha P_n^{(\alpha)}(x) = \mathrm{D}^n(1-x^2)^{n+\alpha},$$

where D stands for $\mathrm{d}/\mathrm{d}x$. Then derive two expressions for $\psi'_{n+1}(x)$:

$$\begin{aligned}
\psi'_{n+1}(x) &= \mathrm{D}^{n+2}(1-x^2)^{n+\alpha+1} \\
&= \mathrm{D}^{n+2}[(1-x^2)(1-x^2)^{n+\alpha}], \\
\psi'_{n+1}(x) &= \mathrm{D}^{n+1}\mathrm{D}(1-x^2)^{n+\alpha+1} \\
&= -2(n+\alpha+1)\mathrm{D}^{n+1}[x(1-x^2)^{n+\alpha}],
\end{aligned}$$

differentiate the two products by applying Leibniz's theorem, and equate the results. This should give a second-order differential equation for $\psi_n(x)$, from which the result follows.]

14. Determine which of the following four systems of $n+1$ weighted polynomials,

$$\{T_i(x)\}, \quad \{\sqrt{1-x^2}\,U_i(x)\}, \quad \{\sqrt{1+x}\,V_i(x)\}, \quad \{\sqrt{1-x}\,W_i(x)\}$$

($0 \le i \le n$) is discretely orthogonal with respect to which of the four following summations

$$\sum_{\text{zeros of } T_{n+1}(x)}, \qquad \sum''_{\text{zeros of } (1-x^2)U_{n-1}(x)},$$

$$\sum^*_{\text{zeros of } (1+x)V_n(x)}, \qquad \sum'_{\text{zeros of } (1-x)W_n(x)}.$$

(Pay particular attention to the cases $i = 0$ and $i = n$.) Find the values of $\langle T_i, T_i \rangle$, and similar inner products, in each case, noting that the result may not be the same for all values of i.

15. Using the discrete inner product $\langle u\,,\,v\rangle = \sum_k u(x_k)v(x_k)$, where $\{x_k\}$ are the zeros of $T_3(x)$, determine monic orthogonal polynomials of degrees 0, 1, 2 using the recurrence (4.21), and verify that they are identical to $\{2^{1-i}T_i(x)\}$.

16. If

$$f_n(x) = \sum_{i=1}^{n} c_i T_{i-1}(x), \qquad (*)$$

where

$$c_i = \frac{1}{n} \sum_{k=1}^{n}{}' f(x_k) T_{i-1}(x_k),$$

and x_k are the zeros of $T_n(x)$, show that $f_n(x_k) = f(x_k)$.

What does the formula (*) provide?

What can we say about the convergence in norm of f_n to f as $n \to \infty$?

17. Using the values of a function $f(x)$ at the zeros $\{x_k\}$ of $T_n(x)$, namely

$$x_k = \cos \frac{(2k+1)\pi}{2n} \quad (k = 0, \ldots, n-1),$$

define another form of discrete Chebyshev transform by

$$\hat{f}(x_k) := \sqrt{\tfrac{2}{n}} \sum_{j=0}^{n-1} T_k(x_j) f(x_j) \quad (k = 0, \ldots, n-1).$$

Use discrete orthogonality to deduce that

$$f(x_j) = \sqrt{\tfrac{2}{n}} \sum_{k=0}^{n-1} T_k(x_j) \hat{f}(x_k) \quad (j = 0, \ldots, n-1).$$

[See Canuto et al. (1988, p.503) for a fast computation procedure based on sets of alternate f values.]

18. Using the values of a function $f(x)$ at the positive zeros $\{x_k\}$ of $T_{2n}(x)$, namely

$$x_k = \cos \frac{(2k+1)\pi}{4n} \quad (k = 0, \ldots, n-1),$$

define another (odd) form of discrete Chebyshev transform by

$$\hat{f}(x_k) := \sqrt{\tfrac{2}{n}} \sum_{j=0}^{n-1} T_{2k+1}(x_j) f(x_j) \quad (k = 0, \ldots, n-1).$$

Deduce that

$$f(x_j) = \sqrt{\tfrac{2}{n}} \sum_{k=0}^{n-1} T_{2k}(x_j) \hat{f}(x_k) \quad (j = 0, \ldots, n-1),$$

and that this transform is self-inverse. [See Canuto et al. (1988, p.504) for a fast computation procedure.]

19. Verify the orthogonality properties (4.85).

20. Show that $\{U_n\}$, $\{V_n\}$ and $\{W_n\}$ are orthogonal over the ellipse E_r with respect to the inner product (4.86) and the appropriate weights.

    Evaluate $\langle T_n\,,\, T_n\rangle$, $\langle U_n\,,\, U_n\rangle$, $\langle V_n\,,\, V_n\rangle$ and $\langle W_n\,,\, W_n\rangle$ for this inner product.

21. Prove that if $\mathcal{A}_r$ denotes the linear space of functions that are analytic throughout the domain $\{z : \left|z + \sqrt{z^2 - 1}\right| \leq r\}$ $(r > 1)$, then $\|\cdot\|_2$, as defined by (4.84), has all of the properties of a norm required by Definition 3.1.