Chapter 6

Chebyshev Interpolation

6.1 Polynomial interpolation

One of the simplest ways of obtaining a polynomial approximation of degree n to a given continuous function $f(x)$ on $[-1,1]$ is to interpolate between the values of $f(x)$ at $n+1$ suitably selected distinct points in the interval. For example, to interpolate at

$$x_1, x_2, \ldots, x_{n+1}$$

by the polynomial

$$p_n(x) = c_0 + c_1 x + \cdots + c_n x^n,$$

we require that

$$c_0 + c_1 x_k + \cdots + c_n x_k^n = f(x_k) \quad (k = 1, \ldots, n+1). \tag{6.1}$$

The equations (6.1) are a set of $n+1$ linear equations for the $n+1$ coefficients $c_0, \ldots, c_n$ that define $p_n(x)$.

Whatever the values of $f(x_k)$, the interpolating polynomial $p_n(x)$ exists and is unique, since the determinant of the linear system (6.1) is non-zero. Specifically

$$\det \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n+1} & x_{n+1}^2 & \cdots & x_{n+1}^n \end{pmatrix} = \prod_{i>j} (x_i - x_j) \neq 0.$$

It is generally not only rather time-consuming, but also numerically unstable, to determine $p_n(x)$ by solving (6.1) as it stands, and indeed many more efficient and reliable formulae for interpolation have been devised.

Some interpolation formulae are tailored to equally spaced points x_1, x_2, ..., x_{n+1}, such as those based on finite differences and bearing the names of Newton and Stirling (Atkinson 1989, for example). Surprisingly however, if we have a free choice of interpolation points, it is not necessarily a good idea to choose them equally spaced. An obvious equally-spaced set for the interval $[-1,1]$ is given for each value of n by

$$x_k = -1 + \frac{2k+1}{n+1} \quad (k = 0, \ldots, n); \tag{6.2}$$

these points are spaced a distance $2/(n+1)$ apart, with half spacings of $1/(n+1)$ between the first and last points and the end points of the interval.

(This set would provide equally spaced interpolation on $(-\infty, \infty)$ if $f(x)$ were periodic with period 2.) However, the following example demonstrates that the points (6.2) are not appropriate for all continuous functions $f(x)$ when n becomes large.

Theorem 6.1 (Runge phenomenon) *If x_k are chosen to be the points (6.2) for each $n \geq 0$, then the interpolating polynomial $p_n(x)$ does* not *converge uniformly on $[-1, 1]$ as $n \to \infty$ for the function $f(x) = 1/(1 + 25x^2)$.*

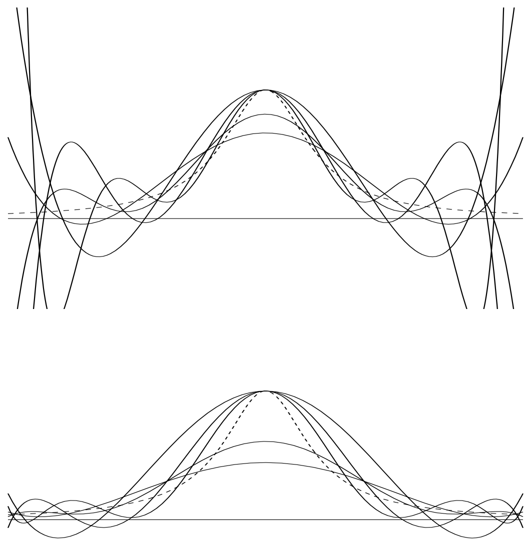

Figure 6.1: Interpolation to $f(x) = 1/(1 + 25x^2)$ by polynomials of degrees 4 to 8 at evenly-spaced points (above) and at Chebyshev polynomial zeros (below)

Proof: We refer the reader to (Mayers 1966) for a full discussion. The function $f(z)$ has complex poles at $z = \pm\frac{1}{5}\mathrm{i}$, which are close to the relevant part of the real axis, and it emerges that such nearby poles are sufficient to prevent uniform convergence. In fact the error $f(x) - p_n(x)$ oscillates wildly close to $x = \pm 1$, for large n. This is illustrated in the upper half of Figure 6.1.

See also (Trefethen & Weideman 1991), where it is noted that Turetskii (1940) showed that the Lebesgue constant for interpolation at evenly-spaced points is asymptotically $2^{n+1}/(\mathrm{e}\, n\, \log n)$. ●●

However, formulae are also available for unequally spaced interpolation, notably Neville's divided-difference algorithm or Aitken's algorithm (Atkinson 1989) and the general formula of Lagrange quoted in Lemma 6.3 below.

A better choice of interpolation points to ensure uniform convergence, though still not necessarily for every continuous function, is the set of zeros of the Chebyshev polynomial $T_{n+1}(x)$, namely (as given in Section 2.2)

$$x = x_k = \cos\frac{(k-\frac{1}{2})\pi}{n+1} \quad (k = 1, \ldots, n+1). \tag{6.3}$$

This choice of points does in fact ensure convergence for the function of Theorem 6.1, and indeed for any continuous $f(x)$ that satisfies a Dini–Lipschitz condition. Thus only a very slight restriction of $f(x)$ is required. This is illustrated in the lower half of Fig. 6.1. See Cheney (1966) or Mason (1982) for a proof of this. We note also from Theorem 6.5 that convergence in a weighted L_2 norm occurs for any continuous $f(x)$.

By expressing the polynomial in terms of Chebyshev polynomials, this choice of interpolation points (6.3) can be made far more efficient and stable from a computational point of view than the equally-spaced set (6.2). So we gain not only from improved convergence but also from efficiency and reliability. We show this in Section 6.3.

Finally, we shall find that we obtain a near-minimax approximation by interpolation at Chebyshev zeros, just as we could by truncating the Chebyshev series expansion — but in this case by a much simpler procedure.

6.2 Orthogonal interpolation

If $\{\phi_i\}$ is any orthogonal polynomial system with ϕ_i of exact degree i then, rather than by going to the trouble of computing an orthogonal polynomial expansion (which requires us to evaluate the inner-product integrals $\langle f\,,\,\phi_i\rangle$), an easier way to form a polynomial approximation $P_n(x)$ of degree n to a given function $f(x)$ is by interpolating $f(x)$ at the $(n+1)$ zeros of $\phi_{n+1}(x)$. In fact, the resulting approximation is often just as good.

The following theorem establishes for general orthogonal polynomials what we already know in the case of Chebyshev polynomials, namely that $\phi_{n+1}(x)$ does indeed have the required $(n+1)$ distinct zeros in the chosen interval.

Theorem 6.2 *If the system $\{\phi_i\}$, with ϕ_i a polynomial of exact degree i, is orthogonal on $[a,b]$ with respect to a non-negative weight $w(x)$, then ϕ_n has exactly n distinct real zeros in $[a,b]$, for every $n \geq 0$.*

Proof: (Snyder 1966, p.7, for example) Suppose that ϕ_n has fewer than n real zeros, or that some of its zeros coincide. Then there are m points $t_1, t_2, \ldots, t_m$ in $[a,b]$, with $0 \leq m < n$, where $\phi_n(x)$ changes sign. Let

$$\Pi_m(x) := \prod_{i=1}^{m}(x - t_i), \quad m \geq 1; \quad \Pi_0(x) := 1.$$

Then Π_m is a polynomial of degree $m < n$, and so must be orthogonal to ϕ_n. But

$$\langle \Pi_m \, , \, \phi_n \rangle = \int_a^b w(x)\Pi_m(x)\phi_n(x)\,\mathrm{d}x \neq 0,$$

since this integrand $w(x)\Pi_m(x)\phi_n(x)$ must have the same sign throughout the interval (except at the m points where it vanishes). We thus arrive at a contradiction. ●●

Since an interpolant samples the values of the function in a discrete set of points only, it is usual to require the function to be in $\mathcal{C}[a,b]$ (i.e., to be continuous), even if we wish to measure the goodness of the approximation in a weaker norm such as $\mathcal{L}_2$.

Some basic facts regarding polynomial interpolation are given by the following lemmas.

Lemma 6.3 *The polynomial of degree n interpolating the continuous function $f(x)$ at the $n+1$ distinct points $x_1, \ldots, x_{n+1}$ can be written as*

$$p_n(x) = \sum_{i=1}^{n+1} \ell_i(x) f(x_i) \tag{6.4}$$

where $\ell_i(x)$ are the usual Lagrange polynomials

$$\ell_i(x) = \prod_{\substack{k=1 \\ k\neq i}}^{n+1} \left(\frac{x - x_k}{x_i - x_k} \right). \tag{6.5}$$

Lemma 6.4 *If $x_1, \ldots, x_{n+1}$ are the zeros of the polynomial $\phi_{n+1}(x)$, then the Lagrange polynomials (6.5) may be written in the form*

$$\ell_i(x) = \frac{\phi_{n+1}(x)}{(x - x_i)\phi'_{n+1}(x_i)}, \tag{6.6}$$

where $\phi'(x)$ denotes the first derivative of $\phi(x)$.

In the special case of the first-kind Chebyshev polynomials, the preceding lemma gives the following specific result.

Corollary 6.4A *For polynomial interpolation at the zeros of the Chebyshev polynomial $T_{n+1}(x)$, the Lagrange polynomials are*

$$\ell_i(x) = \frac{T_{n+1}(x)}{(n+1)(x - x_i)U_n(x_i)},$$

or

$$\begin{aligned} \ell_i(\cos\theta) &= \frac{\cos(n+1)\theta \, \sin\theta_i}{(n+1)(\cos\theta - \cos\theta_i)\sin(n+1)\theta_i} \\ &= -\frac{\sin(n+1)(\theta - \theta_i) \, \sin\theta_i}{(n+1)(\cos\theta - \cos\theta_i)}. \end{aligned} \tag{6.7}$$

The following general result establishes $\mathcal{L}_2$ convergence in this framework of interpolation at orthogonal zeros.

Theorem 6.5 *(Erdös & Turán 1937) If $f(x)$ is in $\mathcal{C}[a,b]$, if $\{\phi_i(x),\ i = 0,1,\ldots\}$ is a system of polynomials (with ϕ_i of exact degree i) orthogonal with respect to $w(x)$ on $[a,b]$ and if $p_n(x)$ interpolates $f(x)$ in the zeros of $\phi_{n+1}(x)$, then*

$$\lim_{n\to\infty} \left(\|f - p_n(x)\|_2\right)^2 = \lim_{n\to\infty} \int_a^b w(x)(f(x) - p_n(x))^2 \,\mathrm{d}x = 0.$$

Proof: The proof is elegant and subtle, and a version for Chebyshev polynomials is given by Rivlin (1974). We give a sketched version.

It is not difficult to show that $\{\ell_i\}$ are orthogonal. By ordering the factors appropriately, we can use (6.6) to write

$$\ell_i(x)\ell_j(x) = \phi_{n+1}(x)\psi_{n-1}(x) \quad (i \neq j)$$

where ψ_{n-1} is a polynomial of degree $n-1$. This must be orthogonal to ϕ_{n+1} and hence

$$\langle \ell_i\,,\, \ell_j\rangle = \langle \phi_{n+1}\,,\, \psi_{n-1}\rangle = 0.$$

Therefore

$$\langle \ell_i\,,\, \ell_j\rangle = 0 \quad (i \neq j). \tag{6.8}$$

Now

$$\|f - p_n\|_2 \le \left\|f - p_n^B\right\|_2 + \left\|p_n^B - p_n\right\|_2$$

where p_n^B is the best $\mathcal{L}_2$ approximation. Therefore, in view of Theorem 4.2, it suffices to prove that

$$\lim_{n\to\infty} \left\|p_n^B - p_n\right\|_2 = 0.$$

Since

$$p_n^B(x) = \sum_{i=1}^{n+1} \ell_i(x) p_n^B(x_i)$$

it follows from (6.4) and (6.8) that

$$\left(\left\|p_n^B - p_n\right\|_2\right)^2 = \sum_{i=1}^{n+1} \langle \ell_i\,,\, \ell_i\rangle \left[f(x_i) - p_n^B(x_i)\right]^2.$$

Provided that $\langle \ell_i\,,\, \ell_i\rangle$ can be shown to be uniformly bounded for all i, the right-hand side of this equality tends to zero by Theorem 4.2. This certainly holds in the case of Chebyshev polynomials, where $\langle \ell_i\,,\, \ell_i\rangle = \dfrac{\pi}{n+1}$. ●●

In the cases $w(x) = (1+x)^{\pm\frac{1}{2}}(1-x)^{\pm\frac{1}{2}}$, Theorem 6.5 gives $\mathcal{L}_2$ convergence properties of polynomial interpolation at Chebyshev polynomial zeros. For example, if x_i are taken to be the zeros of $T_{n+1}(x)$ then

$$\lim_{n\to\infty} \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}}(f(x) - p_n(x))^2 \,\mathrm{d}x = 0.$$

This result can be extended, and indeed Erdös & Feldheim (1936) have established $\mathcal{L}_p$ convergence for all $p > 1$:

$$\lim_{n\to\infty} \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} \left| f(x) - p_n(x) \right|^p \, \mathrm{d}x = 0.$$

In the case of Chebyshev zeros we are able to make more precise comparisons with best approximations (see Section 6.5).

If the function $f(x)$ has an analytic extension into the complex plane, then it may be possible to use the calculus of residues to obtain the following further results.

Lemma 6.6 *If the function $f(x)$ extends to a function $f(z)$ of the complex variable z, which is analytic within a simple closed contour C that encloses the point x and all the zeros $x_1, \ldots, x_{n+1}$ of the polynomial $\phi_{n+1}(x)$, then the polynomial of degree n interpolating $f(x)$ at these zeros can be written as*

$$p_n(x) = \frac{1}{2\pi \mathrm{i}} \oint_C \frac{\{\phi_{n+1}(z) - \phi_{n+1}(x)\} f(z)}{\phi_{n+1}(z)(z-x)} \, \mathrm{d}z \tag{6.9}$$

and its error is

$$f(x) - p_n(x) = \frac{1}{2\pi \mathrm{i}} \oint_C \frac{\phi_{n+1}(x) f(z)}{\phi_{n+1}(z)(z-x)} \, \mathrm{d}z. \tag{6.10}$$

In particular, if $f(x)$ extends to a function analytic within the elliptical contour E_r of Figure 1.5, then we can get a bound on the error of interpolation using the zeros of $T_{n+1}(x)$, implying uniform convergence in this case.

Corollary 6.6A *If the contour C in Lemma 6.6 is the ellipse E_r of (1.34), the locus of the points $\frac{1}{2}(r\mathrm{e}^{\mathrm{i}\theta} + r^{-1}\mathrm{e}^{-\mathrm{i}\theta})$ as θ varies (with $r > 1$, and if $|f(z)| \le M$ at every point z on E_r, then for every real x on $[-1, 1]$ we can show (see Problem 2) from (6.10), using (1.50) and the fact that $|T_{n+1}(x)| \le 1$, that*

$$|f(x) - p_n(x)| \le \frac{(r + r^{-1})M}{(r^{n+1} - r^{-n-1})(r + r^{-1} - 2)}, \quad x \text{ real, } -1 \le x \le 1. \tag{6.11}$$

6.3 Chebyshev interpolation formulae

We showed in Section 4.6 that the Chebyshev polynomials $\{T_i(x)\}$ of degrees up to n are orthogonal in a discrete sense on the set (6.3) of zeros $\{x_k\}$ of $T_{n+1}(x)$. Specifically

$$\sum_{k=1}^{n+1} T_i(x_k) T_j(x_k) = \begin{cases} 0 & i \ne j \ (\le n) \\ n+1 & i = j = 0 \\ \frac{1}{2}(n+1) & 0 < i = j \le n. \end{cases} \tag{6.12}$$

This discrete orthogonality property leads us to a very efficient interpolation formula. Write the nth degree polynomial $p_n(x)$, interpolating $f(x)$ in the points (6.3), as a sum of Chebyshev polynomials in the form

$$p_n(x) = {\sum_{i=0}^{n}}' c_i T_i(x). \tag{6.13}$$

Theorem 6.7 *The coefficients c_i in (6.13) are given by the explicit formula*

$$c_i = \frac{2}{n+1} \sum_{k=1}^{n+1} f(x_k) T_i(x_k). \tag{6.14}$$

Proof: If we set $f(x)$ equal to $p_n(x)$ at the points $\{x_k\}$, then it follows that

$$f(x_k) = {\sum_{i=0}^{n}}' c_i T_i(x_k).$$

Hence, multiplying by $\dfrac{2}{n+1} T_j(x_k)$ and summing,

$$\begin{aligned} \frac{2}{n+1} \sum_{k=1}^{n+1} f(x_k) T_j(x_k) &= {\sum_{i=0}^{n}}' c_i \left\{ \frac{2}{n+1} \sum_{k=1}^{n+1} T_i(x_k) T_j(x_k) \right\} \\ &= c_j, \end{aligned}$$

from (6.12), giving the formula (6.14). ●●

Corollary 6.7A *Formula (6.14) is equivalent to a 'discrete Fourier transform' of the transformed function*

$$g(\theta) = f(\cos\theta).$$

Proof: We have

$$p_n(\cos\theta) = {\sum_{i=0}^{n}}' c_i \cos i\theta$$

with

$$c_i = \frac{2}{n+1} \sum_{k=1}^{n+1} g(\theta_k) \cos i\theta_k, \tag{6.15}$$

with

$$\theta_k = \frac{(k - \frac{1}{2})\pi}{n+1}. \tag{6.16}$$

Thus $\{c_i\}$ are discrete approximations to the true Fourier cosine series coefficients

$$c_i^S = \frac{1}{\pi} \int_{-\pi}^{\pi} g(\theta) \cos i\theta \, \mathrm{d}\theta, \tag{6.17}$$

obtained by applying a trapezoidal quadrature rule to the (periodic) function $g(\theta)$ with equal intervals $\pi/(n+1)$ between the points θ_k. Indeed, a trapezoidal rule approximation to (6.17), valid for any periodic function $g(\theta)$, is

$$c_i^S \simeq \frac{1}{\pi}\frac{\pi}{n+1}\sum_{k=-n}^{n+1} g\left(\frac{(k-\frac{1}{2})\pi}{n+1}\right)\cos\frac{i(k-\frac{1}{2})\pi}{n+1},$$

which gives exactly the formula (6.15) for c_i (when we note that the fact that both $g(\theta)$ and $\cos i\theta$ are even functions implies that the kth and $(1-k)$th terms in the summation are identical). ●●

Thus, Chebyshev interpolation has precisely the same effect as taking the partial sum of an approximate Chebyshev series expansion, obtained by approximating the integrals in the coefficients of the exact expansion by changing the independent variable from x to θ and applying the trapezoidal rule — thus effectively replacing the Fourier transforms c_i^S by discrete Fourier transforms c_i. It is well known among practical mathematicians and engineers that the discrete Fourier transform is a very good substitute for the continuous Fourier transform for periodic functions, and this therefore suggests that Chebyshev interpolation should be a very good substitute for a (truncated) Chebyshev series expansion.

In Sections 4.6.2 and 4.6.3 we obtained analogous discrete orthogonality properties to (6.12), based on the *same* abscissae x_k (zeros of T_{n+1}) but weighted, for the second, third and fourth kind polynomials. However, it is more natural to interpolate a Chebyshev polynomial approximation at the zeros of a polynomial of the same kind, namely the zeros of U_{n+1}, V_{n+1}, W_{n+1} in the case of second, third and fourth kind polynomials. We shall therefore show that analogous discrete orthogonality properties also follow for these new abscissae, and develop corresponding fast interpolation formulae.

6.3.1 Aliasing

We have already seen (Section 6.1) that polynomial interpolation at Chebyshev polynomial zeros is safer than polynomial interpolation at evenly distributed points. Even the former, however, is unreliable if too small a number of points (and so too low a degree of polynomial) is used, in relation to the properties of the function being interpolated.

One mathematical explanation of this remark, particularly as it applies to Chebyshev interpolation, is through the phenomenon of *aliasing*, described as follows.

Suppose that we have a function $f(x)$, having an expansion

$$f(x) = {\sum_{j=0}^{\infty}}' c_j T_j(x) \tag{6.18}$$

in Chebyshev polynomials, which is to be interpolated between its values at the zeros $\{x_k\}$ of $T_{n+1}(x)$ by the finite sum

$$f_n(x) = {\sum_{j=0}^{n}}' \hat{c}_j T_j(x). \tag{6.19}$$

The only information we can use, in order to perform such interpolation, is the set of values of each Chebyshev polynomial at the interpolation points. However, we have the following identity (where $x = \cos\theta$):

$$\begin{aligned} T_j(x) + T_{2n+2\pm j}(x) &= \cos j\theta + \cos(2n+2\pm j)\theta \\ &= \tfrac{1}{2}\cos(n+1)\theta\ \cos(n+1\pm j)\theta \\ &= \tfrac{1}{2}T_{n+1}(x)\, T_{n+1\pm j}(x), \end{aligned} \tag{6.20}$$

so that

$$T_j(x_k) + T_{2n+2\pm j}(x_k) = 0, \quad k = 1, \ldots, n+1. \tag{6.21}$$

Thus $T_{2n+2\pm j}$ is indistinguishable from $-T_j$ over the zeros of T_{n+1}. Figure 6.2 illustrates this in the case $n = 9$, $j = 4$ ($2n+2-j = 16$).

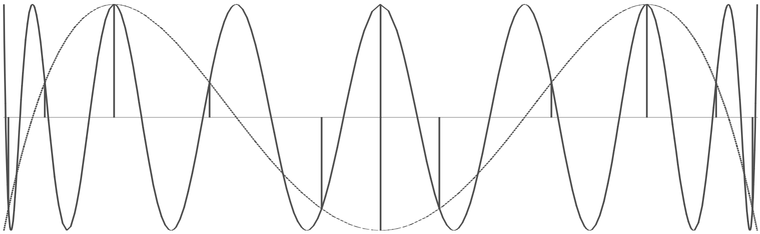

Figure 6.2: $T_{16}(x) = -T_4(x)$ at zeros of $T_{10}(x)$

In consequence, we can say that $f_n(x)$ as in (6.19) interpolates $f(x)$ as in (6.18) between the zeros of $T_{n+1}(x)$ when

$$\hat{c}_j = c_j - c_{2n+2-j} - c_{2n+2+j} + c_{4n+4-j} + c_{4n+4+j} - \cdots, \quad j = 0, \ldots, n. \tag{6.22}$$

(Note that the coefficients c_{n+1}, c_{3n+3}, ...do not figure in (6.22), as they correspond to terms in the expansion that vanish at every interpolation point.)

In effect, the process of interpolation removes certain terms of the expansion (6.18) entirely, while replacing the Chebyshev polynomial in each term after that in $T_n(x)$ by ($\pm 1\times$) a Chebyshev polynomial (its '*alias*') of lower degree. Since the coefficients $\{c_j\}$ tend rapidly to zero for well-behaved functions, the difference between c_j and $\hat{c}_j$ will therefore usually be small, but only if n is taken large enough for the function concerned.

Aliasing can cause problems to the unwary, for instance in working with nonlinear equations. Suppose, for instance that one has a differential equation involving $f(x)$ and $f(x)^3$, and one represents the (unknown) function $f(x)$ in the form $\sum_{j=0}^{\prime n} \hat{c}_j T_j(x)$ as in (6.19). Then one might be tempted to collocate the equation at the zeros of $T_{n+1}(x)$ — effectively carrying out a polynomial interpolation between these points. Instances such as the following, however, cast doubt on the wisdom of this.

In Figure 6.3 we have taken $n = 4$, and show the effect of interpolating the function $T_3(x)^3$ at the zeros of $T_5(x)$. (The expansion of $f_n(x)^3$ includes other products of three Chebyshev polynomials, of course, but this term will suffice.) Clearly the interpolation is poor, the reason being that

$$T_3(x)^3 = \tfrac{1}{4}\left(T_9(x) + 3T_3(x)\right),$$

which aliases to

$$\tfrac{1}{4}\left(-T_1(x) + 3T_3(x)\right).$$

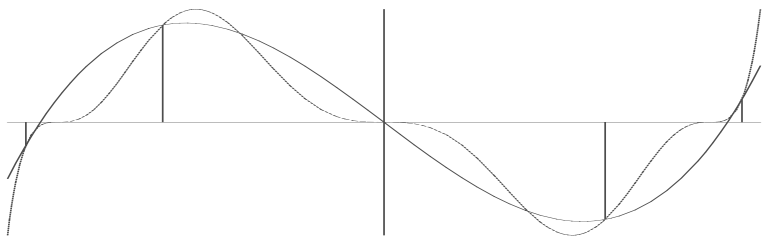

Figure 6.3: $T_3(x)^3$ interpolated at zeros of $T_5(x)$

In contrast, if we could have taken $n = 9$, we could have interpolated $T_3(x)^3$ exactly as shown in Figure 6.4. However, we should then have had

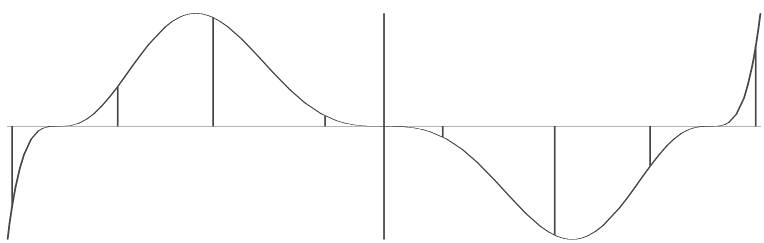

Figure 6.4: $T_3(x)^3$ interpolated (identically) at zeros of $T_{10}(x)$

to consider the effect of aliasing on further products of polynomials of higher order, such as those illustrated in Figures 6.5 and 6.6. There are ways of circumventing such difficulties, which we shall not discuss here.

Much use has been made of the concept of aliasing in estimating quadrature errors (see Section 8.4, where interpolation points and basis functions other than those discussed above are also considered).

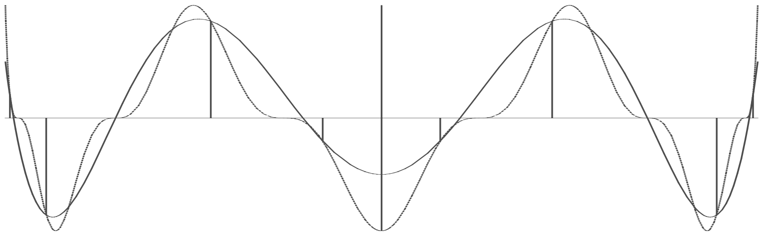

Figure 6.5: $T_6(x)^3$ interpolated at zeros of $T_{10}(x)$

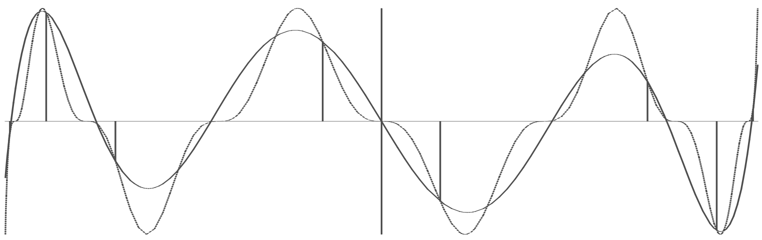

Figure 6.6: $T_7(x)^3$ interpolated at zeros of $T_{10}(x)$

6.3.2 Second-kind interpolation

Consider in this case interpolation by a weighted polynomial $\sqrt{1-x^2}\,p_n(x)$ on the zeros of $U_{n+1}(x)$, namely

$$y_k = \cos\frac{k\pi}{n+2} \quad (k = 1, \ldots, n+1).$$

Theorem 6.8 *The weighted interpolation polynomial to $f(x)$ is given by*

$$\sqrt{1-x^2}\,p_n(x) = \sqrt{1-x^2}\sum_{i=0}^{n} c_i U_i(x) \tag{6.23}$$

with coefficients given by

$$c_i = \frac{2}{n+1}\sum_{k=1}^{n+1}\sqrt{1-y_k^2} f(y_k) U_i(y_k). \tag{6.24}$$

Proof: From (4.50), with $n-1$ replaced by $n+1$,

$$\sum_{k=1}^{n+1}(1-y_k^2)U_i(y_k)U_j(y_k) = \left\{\begin{array}{ll} 0, & i \neq j\ (\leq n); \\ \frac{1}{2}(n+1), & i = j \leq n. \end{array}\right. \tag{6.25}$$

If we set $\sqrt{1-y_k^2} p_n(y_k)$ equal to $f(y_k)$, we obtain

$$f(y_k) = \sqrt{1-y_k^2}\sum_{i=0}^{n} c_i U_i(y_k),$$

and hence, multiplying by $\dfrac{2}{n+1}\sqrt{1-y_k^2}\,U_j(y_k)$ and summing over k,

$$\begin{aligned}\frac{2}{n+1}\sum_{k=1}^{n+1}\sqrt{1-y_k^2}f(y_k)U_j(y_k) &= \sum_{i=0}^{n} c_i\left\{\frac{2}{n+1}\sum_{k=1}^{n+1}(1-y_k^2)U_i(y_k)U_j(y_k)\right\}\\ &= c_i\end{aligned}$$

by (6.25). ●●

Alternatively, we may want to interpolate at the zeros of $U_{n-1}(x)$ together with the points $x = \pm 1$, namely

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

In this case, however, we must express the interpolating polynomial as a sum of first-kind polynomials, when we can use the discrete orthogonality formula (4.45)

$$\sum_{k=0}^{n}{}'' T_i(y_k)T_j(y_k) = \begin{cases} 0, & i \neq j\ (\leq n);\\ \tfrac{1}{2}n, & 0 < i = j < n;\\ n, & i = j = 0;\ i = j = n.\end{cases} \tag{6.26}$$

(Note the double prime indicating that the first and last terms of the sum are to be halved.)

The interpolating polynomial is then

$$p_n(x) = \sum_{i=0}^{n}{}'' c_i T_i(x) \tag{6.27}$$

with coefficients given by

$$c_i = \frac{2}{n}\sum_{k=0}^{n}{}'' f(y_k)T_i(y_k). \tag{6.28}$$

Apart from a factor of $\sqrt{2/n}$, these coefficients make up the discrete Chebyshev transform of Section 4.7.

6.3.3 Third- and fourth-kind interpolation

Taking as interpolation points the zeros of $V_{n+1}(x)$, namely

$$x_k = \cos\frac{(k-\frac{1}{2})\pi}{n+\frac{3}{2}} \quad (k = 1, \ldots, n+1),$$

we have the orthogonality formula, for $i, j \leq n$,

$$\sum_{k=1}^{n+1}(1+x_k)V_i(x_k)V_j(x_k) = \begin{cases} 0 & i \neq j\\ n+\frac{3}{2} & i = j\end{cases} \tag{6.29}$$

(See Problem 14 of Chapter 4.)

Theorem 6.9 *The weighted interpolation polynomial to $\sqrt{1+x}\, f(x)$ is given by*

$$\sqrt{1+x}\, p_n(x) = \sqrt{1+x} \sum_{i=0}^{n} c_i V_i(x) \tag{6.30}$$

where

$$c_i = \frac{1}{n+\frac{3}{2}} \sum_{k=1}^{n+1} \sqrt{1+x_k} f(x_k) V_i(x_k). \tag{6.31}$$

Proof: If we set $\sqrt{1+x_k}\, p_n(x_k)$ equal to $\sqrt{1+x_k} f(x_k)$, we obtain

$$\sqrt{1+x_k}\, f(x_k) = \sqrt{1+x_k} \sum_{i=0}^{n} c_i V_i(x_k),$$

and hence, multiplying by $\dfrac{1}{n+\frac{3}{2}} \sqrt{1+x_k}\, V_j(x_k)$ and summing over k,

$$\begin{aligned}
\frac{1}{n+\frac{3}{2}} \sum_{k=1}^{n+1} (1+x_k) f(x_k) V_j(x_k) &= \sum_{i=0}^{n} c_i \left\{ \frac{1}{n+\frac{3}{2}} \sum_{k=1}^{n+1} (1+x_k) V_i(x_k) V_j(x_k) \right\} \\
&= c_i
\end{aligned}$$

by (6.29). ●●

The same goes for interpolation at the zeros of $W_{n+1}(x)$, namely

$$x_k = \cos \frac{(n-k+2)\pi}{n+\frac{3}{2}} \quad (k = 1, \ldots, n+1),$$

if we replace 'V' by 'W' and '$1+x$' by '$1-x$' throughout.

Alternatively, we may interpolate at the zeros of $V_n(x)$ together with one end point $x = -1$; i.e., at the points

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

where we have the discrete orthogonality formulae (the notation $\sum^{*}$ indicating that the *last* term of the summation is to be halved)

$$\sum_{k=1}^{n+1}{}^{*}\, T_i(x_k) T_j(x_k) = \begin{cases} 0 & i \neq j \ (\leq n) \\ n+\frac{1}{2} & i = j = 0 \\ \frac{1}{2}(n+\frac{1}{2}) & 0 < i = j \leq n. \end{cases} \tag{6.32}$$

The interpolating polynomial is then

$$p_n(x) = \sum_{i=0}^{n}{}' \, c_i T_i(x) \tag{6.33}$$

with coefficients given by

$$c_i = \frac{2}{n+\frac{1}{2}} \sum_{k=1}^{n+1}{}^{*} f(x_k) T_i(x_k). \tag{6.34}$$

6.3.4 Conditioning

In practice, one of the main reasons for the use of a Chebyshev polynomial basis is the good conditioning that frequently results. A number of comparisons have been made of the conditioning of calculations involving various polynomial bases, including $\{x^k\}$ and $\{T_k(x)\}$. A paper by Gautschi (1984) gives a particularly effective approach to this topic.

If a Chebyshev basis is adopted, there are usually three gains:

1. The coefficients generally decrease rapidly with the degree n of polynomial;
2. The coefficients converge individually with n;
3. The basis is well conditioned, so that methods such as collocation are well behaved numerically.

6.4 Best $\mathcal{L}_1$ approximation by Chebyshev interpolation

Up to now, we have concentrated on the $\mathcal{L}_\infty$ or minimax norm. However, the $\mathcal{L}_\infty$ norm is not the only norm for which Chebyshev polynomials can be shown to be minimal. Indeed, a minimality property holds, with a suitable weight function of the form $(1-x)^\gamma(1+x)^\delta$, in the $\mathcal{L}_1$ and $\mathcal{L}_2$ norms, and more generally in the $\mathcal{L}_p$ norm, where p is equal to 1 or an even integer, and this is true for all four kinds of Chebyshev polynomials. Here we look at minimality in the $\mathcal{L}_1$ norm.

The $\mathcal{L}_1$ norm (weighted by $w(x)$) of a function $f(x)$ on $[-1,1]$ is

$$\|f\|_1 := \int_{-1}^{1} w(x)\,|f(x)|\,\mathrm{d}x \tag{6.35}$$

and the Chebyshev polynomials have the following minimality properties in $\mathcal{L}_1$.

Theorem 6.10 $2^{1-n}T_n(x)$ *($n>0$), $2^{-n}U_n(x)$, $2^{-n}V_n(x)$, $2^{-n}W_n(x)$ are the monic polynomials of minimal $\mathcal{L}_1$ norm with respect to the respective weight functions*

$$w(x) = \frac{1}{\sqrt{1-x^2}},\ 1,\ \frac{1}{\sqrt{1-x}},\ \frac{1}{\sqrt{1+x}}. \tag{6.36}$$

Theorem 6.11 *The polynomial $p_{n-1}(x)$ of degree $n-1$ is a best $\mathcal{L}_1$ approximation to a given continuous function $f(x)$ with one of the four weights $w(x)$ given by (6.36) if $f(x) - p_{n-1}(x)$ vanishes at the n zeros of $T_n(x)$, $U_n(x)$, $V_n(x)$, $W_n(x)$, respectively, and at no other interior points of $[-1, 1]$.*

(Note that the condition is sufficient but not necessary.)

Clearly Theorem 6.10 is a special case of Theorem 6.11 (with $f(x) = x^n$), and so it suffices to prove the latter. We first state a classical lemma on the characterisation of best $\mathcal{L}_1$ approximations (Rice 1964, Section 4–4).

Lemma 6.12 *If $f(x) - p_{n-1}(x)$ does not vanish on a set of positive measure (e.g., over the whole of a finite subinterval), where p_{n-1} is a polynomial of degree $n-1$ in x, then p_{n-1} is a best weighted $\mathcal{L}_1$ approximation to f on $[-1, 1]$ if and only if*

$$I_n^{(r)} := \int_{-1}^{1} w(x)\ \mathrm{sgn}[f(x) - p_{n-1}(x)]\ \phi_r(x)\,\mathrm{d}x = 0 \tag{6.37}$$

for $r = 0, 1, \ldots, n-1$, where each $\phi_r(x)$ is any given polynomial of exact degree r.

Using this lemma, we can now establish the theorems.

Proof: (of Theorem 6.11 and hence of Theorem 6.10)

Clearly $\mathrm{sgn}(f(x) - p_{n-1}(x)) = \mathrm{sgn}\, P_n(x)$, where $P_r \equiv T_r,\ U_r,\ V_r,\ W_r$, respectively ($r = 0, 1, \ldots, n$).

Then, taking $\phi_r(x) = P_r(x)$ in (6.37) and making the usual change of variable,

$$I_n^{(r)} = \begin{cases} \displaystyle\int_0^{\pi} \mathrm{sgn}(\cos n\theta)\ \cos r\theta\,\mathrm{d}\theta, \\ \displaystyle\int_0^{\pi} \mathrm{sgn}(\sin(n+1)\theta)\ \sin(r+1)\theta\,\mathrm{d}\theta, \\ \displaystyle\int_0^{\pi} \mathrm{sgn}(\cos(n+\tfrac{1}{2})\theta)\ \cos(r+\tfrac{1}{2})\theta\,\mathrm{d}\theta, \\ \displaystyle\int_0^{\pi} \mathrm{sgn}(\sin(n+\tfrac{1}{2})\theta)\ \sin(r+\tfrac{1}{2})\theta\,\mathrm{d}\theta, \end{cases}$$

respectively. The proof that $I_n^{(r)} = 0$ is somewhat similar in each of these four cases.

Consider the first case. Here, since the zeros of $\cos n\theta$ occur at $(k - \frac{1}{2})\pi/n$ for $k = 1, \ldots, n$, we have

$$\begin{aligned} I_n^{(r)} &= \int_0^{\pi/2n} \cos r\theta\,\mathrm{d}\theta + \sum_{k=1}^{n-1} (-1)^k \int_{(k-\frac{1}{2})\pi/n}^{(k+\frac{1}{2})\pi/n} \cos r\theta\,\mathrm{d}\theta + (-1)^n \int_{(n-\frac{1}{2})\pi/n}^{\pi} \cos r\theta\,\mathrm{d}\theta \\ &= \frac{1}{r}\sin\frac{r\pi}{2n} + \sum_{k=1}^{n-1} (-1)^k \frac{1}{r}\left[\sin\frac{(k+\frac{1}{2})r\pi}{n} - \sin\frac{(k-\frac{1}{2})r\pi}{n}\right] + \end{aligned}$$

$$\begin{aligned}
&\quad + (-1)^{n-1}\frac{1}{r}\sin\frac{(n-\frac{1}{2})r\pi}{n} \\
&= \frac{2}{r}\left[\sin\frac{r\pi}{2n} - \sin\frac{3r\pi}{2n} + \cdots + (-1)^{n-1}\sin\frac{(2n-1)r\pi}{2n}\right] \\
&= \frac{1}{r}\left[\sin\frac{r\pi}{n} - \left\{\sin\frac{r\pi}{n} + \sin\frac{2r\pi}{n}\right\} + \cdots + (-1)^{n-1}\sin\frac{(n-1)r\pi}{n}\right] \Big/ \cos\frac{r\pi}{2n} \\
&= 0.
\end{aligned}$$

We can likewise show that $I_n^{(r)} = 0$ in each of the three remaining cases. Theorems 6.11 and 6.10 then follow very easily from Lemma 6.12 ●●

It follows (replacing n by $n+1$) that the nth degree polynomial $p_n(x)$ interpolating a function $f(x)$ at the zeros of one of the Chebyshev polynomials $T_{n+1}(x)$, $U_{n+1}(x)$, $V_{n+1}(x)$ or $W_{n+1}(x)$, which we showed how to construct in Section 6.3, will in many cases give a best weighted $\mathcal{L}_1$ approximation — subject only to the condition (which we cannot usually verify until after carrying out the interpolation) that $f(x) - p_n(x)$ vanishes nowhere else in the interval.

6.5 Near-minimax approximation by Chebyshev interpolation

Consider a continuous function $f(x)$ and denote the (first-kind) Chebyshev interpolation mapping by J_n. Then

$$(J_n f)(x) = \sum_{k=1}^{n+1} f(x_k)\ell_k(x), \tag{6.38}$$

by the Lagrange formula, and clearly J_n must be a projection, since (6.38) is linear in f and exact when f is a polynomial of degree n. From Lemma 5.13, J_n is near-minimax within a relative distance $\|J_n\|_\infty$.

Now

$$|(J_n f)(x)| \le \sum_{k=1}^{n+1} \|f\|_\infty |\ell_k(x)| .$$

Hence

$$\begin{aligned}
\|J_n\|_\infty &= \sup_f \frac{\|J_n f\|_\infty}{\|f\|_\infty} \\
&= \sup_f \sup_{x\in[-1,1]} \frac{|(J_n f)(x)|}{\|f\|_\infty} \\
&\le \sup_f \sup_{x\in[-1,1]} \sum_{k=0}^{n} |\ell_k(x)| \\
&= \mu_n
\end{aligned} \tag{6.39}$$

where

$$\mu_n = \sup_{x\in[-1,1]} \sum_{k=1}^{n+1} |\ell_k(x)| \,. \qquad (6.40)$$

Now if $\sum_{k=0}^{n} |\ell_k(x)|$ attains its extremum at $x = \xi$, we can define a continuous function $\phi(x)$ such that

$$\begin{aligned} \|\phi\|_\infty &\le 1, \\ \phi(x_k) &= \operatorname{sgn}(\ell_k(\xi)). \end{aligned}$$

Then, from (6.38),

$$(J_n\phi)(\xi) = \sum_{k=1}^{n+1} |\ell_k(\xi)| = \mu_n,$$

whence

$$\|J_n\|_\infty \ge \|J_n\phi\|_\infty \ge \mu_n. \qquad (6.41)$$

Inequalities (6.39) and (6.41) together give us

$$\|J_n\|_\infty = \mu_n.$$

What we have written so far applies to any Lagrange interpolation operator. If we specialise to first-kind Chebyshev interpolation, where $\ell_k(x)$ is as given by Corollary 6.4A, then we have the following asymptotic bound on $\|J_n\|_\infty$.

Theorem 6.13 *If $\{x_k\}$ are the zeros of $T_{n+1}(x)$, then*

1.

$$\mu_n = \frac{1}{\pi} \sum_{k=1}^{n+1} \left| \cot \frac{(k-\frac{1}{2})\pi}{2(n+1)} \right|,$$

2.

$$\mu_n = \frac{2}{\pi} \log n + 0.9625 + O(1/n) \ \text{ as } n \to \infty.$$

Proof: For the details of the proof, the reader is referred to Powell (1967) or Rivlin (1974). See also Brutman (1978). ●●

The following classical lemma then enables us to deduce convergence properties.

Lemma 6.14 (Jackson's theorem) *If $\omega(\delta)$ is the modulus of continuity of $f(x)$, then the minimax polynomial approximation $B_n f$ of degree n to f satisfies*

$$\|f - B_n f\|_\infty \le \omega(1/n).$$

Corollary 6.14A *If $(J_n f)(x)$ interpolates $f(x)$ in the zeros of $T_{n+1}(x)$, and if $f(x)$ is Dini–Lipschitz continuous, then $(J_n f)(x)$ converges uniformly to $f(x)$ as $n \to \infty$.*

Proof: By the definition of Dini–Lipschitz continuity, $\omega(\delta)\log\delta \to 0$ as $\delta \to 0$. By Theorem 5.12

$$\begin{aligned}
\|f - J_n f\|_\infty &\leq (1 + \|J_n\|)_\infty \|f - B_n f\|_\infty \\
&\leq (1 + \mu_n)\omega(1/n) \\
&\sim \frac{2}{\pi}\omega(1/n)\log n \\
&= -\frac{2}{\pi}\omega(\delta)\log\delta \quad (\delta = 1/n) \\
&\to 0 \text{ as } \delta \to 0; \text{ i.e., as } n \to \infty. \quad ●●
\end{aligned}$$

In closing this chapter, we remind the reader that further interpolation results have been given earlier in Chapter 4 in the context of orthogonality. See in particular Sections 4.3.2 and 6.2.

6.6 Problems for Chapter 6

1. Prove Lemmas 6.3 and 6.4, and deduce Corollary 6.4A.

2. Prove Corollary 6.6A.

3. Find expressions for the coefficients (6.14) of the nth degree interpolating polynomial when $f(x) = \operatorname{sgn} x$ and $f(x) = |x|$, and compare these with the coefficients in the Chebyshev expansions (5.11) and (5.12).

4. List the possibilities of aliasing in the following interpolation situations:

   (a) Polynomials U_j of the second kind on the zeros of $T_{n+1}(x)$,

   (b) Polynomials V_j of the third kind on the zeros of $T_{n+1}(x)$,

   (c) Polynomials U_j on the zeros of $(1 - x^2)U_{n-1}(x)$,

   (d) Polynomials T_j on the zeros of $(1 - x^2)U_{n-1}(x)$,

   (e) Polynomials V_j on the zeros of $(1 - x^2)U_{n-1}(x)$,

   (f) Polynomials U_j on the zeros of $U_{n+1}(x)$,

   (g) Polynomials T_j on the zeros of $U_{n+1}(x)$.

5. Give a proof of Theorem 6.11 for the case of the function U_r.

6. Using Theorem 6.11, consider the lacunary series partial sum

$$f_n(x) = \sum_{k=1}^{n} c_k U_{2^k-1}(x).$$

Assuming that the series is convergent to $f = \lim_{n\to\infty} f_n$, show that $f - f_n$, for instance, vanishes at the zeros of U_{2^n-1}. Give sufficient conditions for f_n to be a best $\mathcal{L}_1$ approximation to f for every n.

7. Show that the $n+1$ zeros of $T_{n+1}(z) - T_{n+1}(z^*)$ are distinct and lie on E_r, for a suitable fixed point z^* on E_r $(r > 1)$. Fixing r, find the zeros for the following choices of z^*:

   (a) $z^* = \frac{1}{2}(r + r^{-1})$,
   (b) $z^* = -\frac{1}{2}(r + r^{-1})$,
   (c) $z^* = \frac{1}{2}\mathrm{i}(r - r^{-1})$,
   (d) $z^* = -\frac{1}{2}\mathrm{i}(r - r^{-1})$.

8. If $f_n(z)$ is a polynomial of degree n interpolating $f(z)$, continuous on the ellipse E_r and analytic in its interior, find a set of interpolation points z_k $(k = 1, \ldots, n+1)$ on E_r such that

   (a) f_n is near-minimax within a computable relative distance σ_n on E_r, giving a formula for σ_n;

   (b) this result is valid as $r \to 1$; i.e., as the ellipse collapses to the line segment $[-1, 1]$.

   To effect (b), show that it is necessary to choose the interpolation points asymmetrically across the x-axis, so that points do not coalesce.