4

Polynomial interpolation

4.1 The Lagrange interpolation formula

If one decides to approximate a function $f \in \mathscr{C}[a, b]$ by a polynomial

$$p(x) = \sum_{i=0}^{n} c_i x^i, \qquad a \leqslant x \leqslant b, \tag{4.1}$$

one has the problem of specifying the coefficients $\{c_i; i = 0, 1, \ldots, n\}$. The most straightforward method is to calculate the value of f at $(n+1)$ distinct points $\{x_i; i = 0, 1, \ldots, n\}$ of $[a, b]$, and to satisfy the equations

$$p(x_i) = f(x_i), \qquad i = 0, 1, \ldots, n. \tag{4.2}$$

We note that there are as many conditions as coefficients, and the following theorem shows that they determine $p \in \mathscr{P}_n$ uniquely.

Theorem 4.1

Let $\{x_i; i = 0, 1, \ldots, n\}$ be any set of $(n+1)$ distinct points in $[a, b]$, and let $f \in \mathscr{C}[a, b]$. Then there is exactly one polynomial $p \in \mathscr{P}_n$ that satisfies the equations (4.2).

Proof. For $k = 0, 1, \ldots, n$, let l_k be the function

$$l_k(x) = \prod_{\substack{j=0 \\ j \neq k}}^{n} (x - x_j)/(x_k - x_j), \quad a \leqslant x \leqslant b. \tag{4.3}$$

We note that $l_k \in \mathscr{P}_n$ and that the equations

$$l_k(x_i) = \delta_{ki}, \qquad i = 0, 1, \ldots, n, \tag{4.4}$$

hold, where δ_{ki} has the value

$$\delta_{ki} = \begin{cases} 1, & k = i, \\ 0, & k \neq i. \end{cases} \tag{4.5}$$

It follows that the function

$$p=\sum_{k=0}^{n} f(x_k)\, l_k \tag{4.6}$$

is in $\mathscr{P}_n$ and it satisfies the required interpolation conditions (4.2). To show uniqueness, suppose that the equations (4.2) are satisfied by both $p\in\mathscr{P}_n$ and $q\in\mathscr{P}_n$. Then the difference $(p-q)$ is in $\mathscr{P}_n$ and it has roots at the points $\{x_i;\, i=0, 1, \ldots, n\}$. However, a polynomial of degree at most n that has $(n+1)$ distinct roots is identically zero. Therefore p is equal to q. □

The numerical value of the interpolating polynomial $p(x)$ for any fixed x in $[a, b]$ can be calculated by first computing the numbers (4.3) for $k=0, 1, \ldots, n$, and then by substituting them in the equation

$$p(x)=\sum_{k=0}^{n} f(x_k)\, l_k(x). \tag{4.7}$$

This method is called the Lagrange interpolation formula. There are many other algorithms for calculating $p(x)$ that are equivalent in exact arithmetic. They differ, however, in the accuracy that is obtained in the presence of computer rounding errors, and in the amount of work that is done when they are applied. One of the most successful algorithms, which is called Newton's interpolation method, is described in the next chapter.

The uniqueness property, proved in Theorem 4.1, allows us to regard the interpolation process as an operator from $\mathscr{C}[a, b]$ to $\mathscr{P}_n$, which depends on the choice of the fixed points $\{x_i;\, i=0,\ 1, \ldots, n\}$. The operator is a projection because, if $f\in\mathscr{P}_n$, then we may satisfy the interpolation conditions (4.2) by making p equal to f. Moreover, because the functions l_k $(k=0, 1, \ldots, n)$ are independent of f, equation (4.6) shows that the operator is linear. Therefore we may apply Theorem 3.1, and we find in Section 4.4 that it gives some interesting results.

When the function values $\{f(x_i);\, i=0, 1, \ldots, n\}$ cannot be obtained exactly, it may be important to know the contribution that their errors make to the calculated polynomial p. Equation (4.6) answers this question directly, for, if the true function value $f(x_k)$ is replaced by the approximation $\{f(x_k)+\varepsilon_k\}$ for $k=0, 1, \ldots, n$, we see that the change to p is the expression $\sum \varepsilon_k l_k$.

The Lagrange interpolation formula provides some algebraic relations that are useful in later work. They come from our remark that the interpolation process is a projection operator. In particular, for $0\leqslant i\leqslant n$, we let f be the function

$$f(x)=x^i, \qquad a\leqslant x\leqslant b, \tag{4.8}$$

in order to obtain from expression (4.7) the equation

$$\sum_{k=0}^{n} x_k^i l_k(x) = x^i, \qquad a \leqslant x \leqslant b. \tag{4.9}$$

The value $i=0$ gives the identity

$$\sum_{k=0}^{n} l_k(x) = 1, \qquad a \leqslant x \leqslant b, \tag{4.10}$$

which is useful for checking the numbers $\{l_k(x);\ k=0, 1, \ldots, n\}$ when the Lagrange interpolation method is applied. Moreover, by substituting the definition (4.3) in equation (4.9), and then by considering the coefficient of x^n, we find the identity

$$\sum_{k=0}^{n} \frac{x_k^i}{\prod\limits_{\substack{j=0\\ j\neq k}}^{n} (x_k - x_j)} = \delta_{in}, \qquad i=0, 1, \ldots, n. \tag{4.11}$$

4.2 The error in polynomial interpolation

We use the notation e for the error function of an approximation, and in this chapter it has the value

$$e(x) = f(x) - p(x), \qquad a \leqslant x \leqslant b, \tag{4.12}$$

where p is the polynomial in $\mathscr{P}_n$ that satisfies the interpolation conditions (4.2). It should be clear that, if we change f by adding to it an element of $\mathscr{P}_n$, then the interpolation process automatically adds the same element to p, which leaves e unchanged. Expressions for the error should show this property. It is therefore appropriate, when $f \in \mathscr{C}^{(n+1)}[a, b]$, to state e in terms of the derivative $f^{(n+1)}$, which is done in our next theorem.

Theorem 4.2

For any set of distinct interpolation points $\{x_i;\ i=0, 1, \ldots, n\}$ in $[a, b]$ and for any $f \in \mathscr{C}^{(n+1)}[a, b]$, let p be the element of $\mathscr{P}_n$ that satisfies the equations (4.2). Then, for any x in $[a, b]$, the error (4.12) has the value

$$e(x) = \frac{1}{(n+1)!} \prod_{j=0}^{n} (x - x_j) f^{(n+1)}(\xi), \tag{4.13}$$

where ξ is a point of $[a, b]$ that depends on x.

Proof. Two methods are used in this book to express errors in terms of derivatives. One is to apply the Taylor series expansion, and the other one is to use Rolle's theorem several times. Rolle's theorem states that, if a continuously differentiable function is zero at two points, then its derivative is zero at an intermediate point. By using this result inductively, we deduce that, if a function $g \in \mathscr{C}^{(n+1)}[a, b]$ is zero at $(n+2)$ distinct points of $[a, b]$, then its $(n+1)$th derivative has at least one zero in $[a, b]$. The present proof depends on this fact.

We note first that, if x is in the point set $\{x_i; i=0, 1, \ldots, n\}$, then equation (4.13) holds, because both sides of the equation are equal to zero. Otherwise we define the function g by the equation

$$g(t)=f(t)-p(t)-e(x)\prod_{i=0}^{n}\frac{(t-x_i)}{(x-x_i)}, \qquad a \leqslant t \leqslant b, \tag{4.14}$$

and it is important to note that t is the variable, the value of x being fixed. We see that $g \in \mathscr{C}^{(n+1)}[a, b]$, and that $g(t)$ is zero both when $t=x$ and when t is in the point set $\{x_i; i=0, 1, \ldots, n\}$. Therefore there exists a point ξ in $[a, b]$ at which the equation

$$g^{(n+1)}(\xi)=0 \tag{4.15}$$

is satisfied. By substituting the definition (4.14) in this equation, and by rearranging terms, we find the required result (4.13). □

A helpful way of remembering this result is to let f be the function

$$f(x)=x^{n+1}, \qquad a \leqslant x \leqslant b. \tag{4.16}$$

In this case the error function is the polynomial

$$e(x)=x^{n+1}-p(x), \qquad a \leqslant x \leqslant b, \tag{4.17}$$

and, because the error is zero at the interpolation points $\{x_i; i= 0, 1, \ldots, n\}$, $e(x)$ must be a multiple of the product

$$\prod_{j=0}^{n}(x-x_j). \tag{4.18}$$

The multiplying factor is the term $f^{(n+1)}(\xi)$ times a constant, which has to have the value $1/(n+1)!$, in order that the coefficient of x^{n+1} in $e(x)$ is equal to one, as required by equation (4.17).

Some applications of Theorem 4.2 are as follows. If a bound on $\|f^{(n+1)}\|_\infty$ is known, then expression (4.13) gives a bound on the error of polynomial interpolation. Similarly, an estimate of the term $f^{(n+1)}(\xi)$ provides an estimate of the interpolation error, which is discussed further in the next chapter. Moreover, Theorem 4.2 is useful sometimes when one wishes to compare polynomial interpolation with some other linear approximation operator that is exact for $f \in \mathscr{P}_n$. If the error of the alternative operator is expressed in terms of $f^{(n+1)}$, then equation (4.13) helps to show which approximation method is more accurate.

4.3 The Chebyshev interpolation points

This section concerns the choice of the interpolation points $\{x_i;\ i=0,1,\ldots,n\}$. Most of the conclusions are obtained by applying polynomial interpolation to a particular function f, known as Runge's example. It is the function

$$f(x)=1/(1+x^2), \qquad -5\leq x\leq 5. \tag{4.19}$$

Because most of the variation in f occurs in the middle of the range $-5\leq x\leq 5$, the discussion given in Section 3.3 shows that it is not really suitable to approximate f by a single polynomial. We have to choose a polynomial of very high degree if we wish to achieve high accuracy. Therefore the example serves quite well to show the kinds of difficulty that can occur in polynomial interpolation. In particular, we find that the positions of the interpolation points $\{x_i;\ i=0,1,\ldots,n\}$ are important when n is large.

If the interpolation points are spaced uniformly

$$x_i=-5+10i/n, \qquad i=0,1,\ldots,n, \tag{4.20}$$

then the size of the error function (4.12) near the ends of the range $-5\leq x\leq 5$ is interesting. We let $x_{n-\frac{1}{2}}$ be the point

$$x_{n-\frac{1}{2}}=5-5/n, \tag{4.21}$$

which is the mid-point of the last interval between interpolation points. The value of $p(x_{n-\frac{1}{2}})$ was found by Lagrange interpolation for $n=2,4,\ldots,20$, and the results are shown in Table 4.1. We see that the error almost doubles in magnitude each time n is increased by two. Therefore it

Table 4.1. *The dependence of $e(x_{n-\frac{1}{2}})$ on n in Runge's example*

n	$f(x_{n-\frac{1}{2}})$	$p(x_{n-\frac{1}{2}})$	$e(x_{n-\frac{1}{2}})$
2	0.137 931	0.759 615	−0.621 684
4	0.066 390	−0.356 826	0.423 216
6	0.054 463	0.607 879	−0.553 416
8	0.049 651	−0.831 017	0.880 668
10	0.047 059	1.578 721	−1.531 662
12	0.045 440	−2.755 000	2.800 440
14	0.044 334	5.332 743	−5.288 409
16	0.043 530	−10.173 867	10.217 397
18	0.042 920	20.123 671	−20.080 751
20	0.042 440	−39.952 449	39.994 889

is futile to try to improve the accuracy of the approximation by increasing the value of n.

The reason for the large values of $e(x)$ shown in Table 4.1 can be found from the form of the error function when $n=20$. Values of this function are given in Table 4.2 at the points that are midway between the interpolation points in $0 \leqslant x \leqslant 5$. Negative values of x are omitted because f and p are both even functions of x. The function (4.18), which is called prod(x), is also tabulated. The most important feature of the table is that the very rapid increase in the tabulated values of $e(x)$ also occurs in the tabulated values of prod(x). Indeed the ratio $e(x)/\text{prod}(x)$ is almost constant.

It follows, therefore, that in this example the dependence on x of the term $f^{(n+1)}(\xi)$ in equation (4.13) does not make much difference to the form of $e(x)$. A good practical strategy is to assume that this property remains true if the positions of the interpolation points $\{x_i; i=0, 1, \ldots, n\}$ are altered. Therefore we wish to find interpolation points that do not give large variations in the heights of the peaks of prod(x). By bunching interpolation points near the ends of the range, the very large peaks of prod(x) can be reduced, at the expense of increasing the heights of the small peaks near the centre of the range $-5 \leqslant x \leqslant 5$. The interpolation points that equalize the peak heights are called the Chebyshev interpolation points, and they are found by making use of 'Chebyshev polynomials'.

For the range $-1 \leqslant x \leqslant 1$, the Chebyshev polynomial of degree n is the function T_n that satisfies the equation

$$T_n(\cos\theta) = \cos(n\theta), \tag{4.22}$$

Table 4.2. *An example of equally spaced interpolation points* ($n=20$)

x	$f(x)$	$p(x)$	$e(x)$	prod(x)
0.25	0.941 176	0.942 490	−0.001 314	2.05×10^6
0.75	0.640 000	0.636 755	0.003 245	-2.48×10^6
1.25	0.390 244	0.395 093	−0.004 849	3.64×10^6
1.75	0.246 154	0.238 446	0.007 708	-6.56×10^6
2.25	0.164 948	0.179 763	−0.014 814	1.46×10^7
2.75	0.116 788	0.080 660	0.036 128	-4.12×10^7
3.25	0.086 486	0.202 423	−0.115 936	1.51×10^8
3.75	0.066 390	−0.447 052	0.513442	-7.56×10^8
4.25	0.052 459	3.454 958	−3.402 499	5.59×10^9
4.75	0.042 440	−39.952 449	39.994 889	-7.27×10^{10}

which is equivalent to the equation

$$T_n(x) = \cos(n \cos^{-1} x), \qquad -1 \leqslant x \leqslant 1. \tag{4.23}$$

An easy way of imagining $T_n(x)$ as a function of x is to expand $\cos(n\theta)$ in powers of $\cos\theta$, and to write x in place of $\cos\theta$. Hence $T_n \in \mathscr{P}_n$, and the identity

$$\cos[(n+1)\theta] + \cos[(n-1)\theta] = 2\cos\theta\cos(n\theta) \tag{4.24}$$

gives the recurrence relation

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \qquad -1 \leqslant x \leqslant 1. \tag{4.25}$$

Chebyshev polynomials have many applications in approximation theory, and they are useful now because the heights of the peaks of the function

$$T_n(x) = \cos(n\theta), \qquad x = \cos\theta, \tag{4.26}$$

are all equal to one. We can force prod (x) to be a multiple of $T_{n+1}(x)$ by letting the interpolation points $\{x_i;\, i = 0, 1, \ldots, n\}$ be the roots of the polynomial T_{n+1}, which gives the points

$$x_i = \cos\left\{\frac{[2(n-i)+1]\pi}{2(n+1)}\right\}, \qquad i = 0, 1, \ldots, n. \tag{4.27}$$

In order to adapt these values to a general range $a \leqslant x \leqslant b$, we introduce real parameters λ and μ, and we define the points

$$x_i = \lambda + \mu \cos\left\{\frac{[2(n-i)+1]\pi}{2(n+1)}\right\}, \qquad i = 0, 1, \ldots, n, \tag{4.28}$$

to be Chebyshev interpolation points. By construction they have the property that the magnitudes of the peaks of the polynomial (4.18) are all equal, which helps usually to reduce the greatest value of the error function (4.13), provided that x_0 is close to a and x_n is close to b. We really want to choose the interpolation points in a way that makes the expression

$$\max_{a \leqslant x \leqslant b} |\text{prod}(x)| \tag{4.29}$$

small. A theorem in Chapter 7 shows that this expression is minimized over all sets $\{x_i;\, i = 0, 1, \ldots, n\}$ if λ and μ have the values

$$\left.\begin{aligned} \lambda &= \tfrac{1}{2}(a+b) \\ \mu &= \tfrac{1}{2}(b-a) \end{aligned}\right\} \tag{4.30}$$

in equation (4.28).

In order to show that the use of Chebshev interpolation points can improve on the accuracy that is shown in Table 4.2, we let $\{x_i;\, i = 0, 1, \ldots, n\}$ have the values (4.28), where $n = 20$ and where λ and μ are such that $x_0=-5$ and $x_{20}=5$. The Lagrange interpolation method was applied again to Runge's function (4.19). Table 4.3 shows the errors of interpolation at the positive values of x where $|\text{prod}(x)|$ is greatest. We find that the greatest value of $|e(x)|$ is smaller than in Table 4.2 by a factor of over two thousand, and the cost of this gain is that the small errors near the centre of the range $-5 \leqslant x \leqslant 5$ are increased by about a factor of five. Now all the variations in the tabulated values of $e(x)$ are due to the term $f^{(n+1)}(\xi)$ in equation (4.13).

It is also of interest to note the improvement over Table 4.1 that can be obtained by using Chebyshev interpolation points. Therefore, for $n=2$, $4, \ldots, 20$, we let the set $\{x_i; i=0, 1, \ldots, n\}$ be defined by equation (4.28), where, as in the last paragraph, the values of λ and μ are such that $x_0=-5$ and $x_n=5$. Thus an interpolating polynomial $p \in \mathscr{P}_n$ is obtained for each n. By applying Lagrange interpolation for several values of x, the

Table 4.3. *An example of Chebyshev interpolation points* ($n=20$)

x	$f(x)$	$p(x)$	$e(x)$
0.374 698	0.876 886	0.887 135	−0.010 249
1.115 724	0.445 466	0.429 963	0.015 503
1.831 827	0.229 590	0.242 708	−0.013 119
2.507 010	0.137 266	0.126 532	0.010 734
3.126 190	0.092 824	0.101 876	−0.009 052
3.675 537	0.068 920	0.061 018	0.007 902
4.142 778	0.055 058	0.062 173	−0.007 115
4.517 476	0.046 712	0.040 130	0.006 582
4.791 261	0.041 743	0.047 981	−0.006 238
4.958 018	0.039 090	0.033 045	0.006 045

Table 4.4. *The maximum error when Chebyshev interpolation points are used*

n	x	$f(x)$	$p(x)$	$e(x)$
2	2.024 604	0.196 116	0.842 345	−0.646 229
4	1.393 399	0.339 765	0.761 908	−0.442 143
6	1.097 876	0.453 447	0.727 637	−0.274 191
8	0.912 455	0.545 680	0.721 700	−0.176 020
10	0.781 995	0.620 534	0.732 455	−0.111 921
12	0.684 167	0.681 159	0.751 878	−0.070 718
14	1.526 988	0.300 148	0.252 887	0.047 260
16	1.356 570	0.352 078	0.319 037	0.033 040
18	1.221 054	0.401 449	0.378 684	0.022 765
20	1.110 623	0.447 731	0.432 224	0.015 507

maximum value of $|e(x)|$ was calculated. The values of x that maximize the error function and the corresponding values of f, p and e are shown in Table 4.4. We see that the use of Chebyshev interpolation points is so much better than equally spaced ones, that now the accuracy of the approximation improves when n is increased.

4.4 The norm of the Lagrange interpolation operator

Theorem 3.1 provides an excellent reason for studying the norm of the Lagrange interpolation operator. We use the ∞-norm for the elements of $\mathscr{C}[a, b]$, we assume that the set of interpolation points $\{x_i; i = 0, 1, \ldots, n\}$ has been chosen and, for each f in $\mathscr{C}[a, b]$, we let $X(f)$ be the element of $\mathscr{P}_n$ that is defined by the conditions (4.2). The value of $\|X\|$ is the subject of our next theorem.

Theorem 4.3

The norm of the Lagrange interpolation operator has the value

$$\|X\| = \max_{a \leq x \leq b} \sum_{k=0}^{n} |l_k(x)|, \tag{4.31}$$

where the functions $\{l_k; k = 0, 1, \ldots, n\}$ are defined by equation (4.3).

Proof. The definition of a norm and equation (4.6) give the identity

$$\begin{aligned} \|X\| &= \sup_{\|f\| \leq 1} \|X(f)\| \\ &= \sup_{\|f\| \leq 1} \max_{a \leq x \leq b} \left| \sum_{k=0}^{n} f(x_k) l_k(x) \right| \\ &= \max_{a \leq x \leq b} \sup_{\|f\| \leq 1} \left| \sum_{k=0}^{n} f(x_k) l_k(x) \right| \\ &= \max_{a \leq x \leq b} \sum_{k=0}^{n} |l_k(x)|, \end{aligned} \tag{4.32}$$

which is the required result. □

We note that the method of proof is to treat the supremum over f in equation (4.32) before the maximum over x. Often expressions for norms are suprema of maxima, and it is usually helpful, especially in the case of interpolation operators, to take account of the conditions on f before maximizing over x.

Theorem 3.1 states that the error $\|f - X(f)\|$ is within the factor $[1 + \|X\|]$ of the least error

$$d^*(f) = \min_{p \in \mathscr{P}_n} \|f - p\| \tag{4.33}$$

that can be achieved by approximating f by a member of $\mathcal{P}_n$. Hence we obtain from Tables 4.2 and 4.4 a lower bound on $\|X\|$, where X is the interpolation operator in the case when $n=20$ and the interpolation points have the equally spaced values (4.20). Because Table 4.4 shows that 0.015 507 is an upper bound on $d^*(f)$, it follows from Theorem 3.1 and Table 4.2 that the inequality

$$\|X\| \geq (39.994\,889/0.015\,507)-1 \tag{4.34}$$

holds. Hence $\|X\|$ is rather large, and in fact it is equal to 10 986.71, which was calculated by evaluating the function on the right-hand side of equation (4.31) for several values of x. Table 4.5 gives $\|X\|$ for $n = 2, 4, \ldots, 20$ for the interpolation points (4.20). It also gives the value of $\|X\|$ for the Chebyshev interpolation points (4.28) that are relevant to Table 4.4, where λ and μ are such that $x_0=-5$ and $x_n=5$.

Table 4.5 shows clearly that, if the choice of interpolation points is independent of f, and if n is large, then it is safer to use Chebyshev points instead of equally spaced ones. Indeed, if $n=20$ and if Chebyshev points are preferred, then it follows from Theorem 3.1 that, for all $f \in \mathscr{C}[-5, 5]$, the maximum error of the interpolating polynomial is within the factor 3.48 of the least maximum error that can be achieved. However, if the interpolation points are equally spaced, then the form of the error function shown in Table 4.2 is typical, where the maximum error is much larger than necessary. Moreover, another good practical reason for keeping $\|X\|$ small is that it makes the calculated polynomial less sensitive to errors in the data.

Table 4.5. *The norms of some interpolation operators*

n	Equally spaced points	Chebyshev points
2	1.25	1.25
4	2.21	1.57
6	4.55	1.78
8	10.95	1.94
10	29.90	2.07
12	89.32	2.17
14	283.21	2.27
16	934.53	2.34
18	3 171.37	2.42
20	10 986.71	2.48

The results in Table 4.5 are not special to the range $-5 \leqslant x \leqslant 5$, because a general linear transformation of the form

$$x \to \alpha x + \beta, \qquad \alpha > 0, \tag{4.35}$$

where α and β are real parameters, which changes $[a, b]$ to $[\alpha a + \beta, \alpha b + \beta]$ and $\{x_i; i = 0, 1, \ldots, n\}$ to $\{\alpha x_i + \beta; i = 0, 1, \ldots, n\}$, does not alter the value of $\|X\|$. The reason is that this transformation just introduces the factor α^n into the numerator and denominator of the definition (4.3) and these factors cancel each other. Hence the transformation stretches or contracts the graphs of l_k $(k = 0, 1, \ldots, n)$ in the x-direction, but it leaves them unaltered in the y-direction. Thus the value of expression (4.31) does not change, and identities like equation (4.10) are preserved.