Chapter 1

Polynomial Interpolation

1.1 Introduction

Suppose that we have a two sets of $n+1$ real numbers $\{x_i\}_{i=0}^{n+1}$ and $\{y_i\}_{i=0}^{n+1}$, and that the x_i are *strictly* increasing: $x_0 < x_1 < x_2 < \cdots < x_n$. *Interpolation* problems are of the form: *Find a function* p*, that is continuous and defined on* $[x_0, x_n]$*, such that*

$$p(x_k) = y_k, \quad \textit{for} \quad k = 0, 1, \ldots, n.$$

We say that p *interpolates* the points (x_0, y_0), (x_1, y_1), ..., (x_n, y_n).

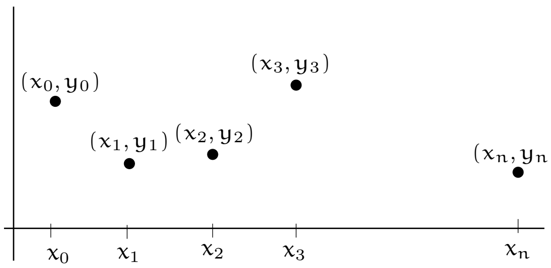


Fig. 1.1: Some points to interpolate

Why Bother? There are several possibilities, including

- The points belong to an underlying, but unknown function. We wish to establish likely values of f at points other than x_0, x_1, ..., x_n. The values of f may have been obtained from physical experiments, or numerical procedures (e.g., Newton's method for initial value problems). Or it may be that some values of the function are easily available. For example $2! = 2$, and $3! = 6$, but what about $2\frac{1}{2}!$ or $\pi!$?
- We may know the function, but prefer to work with an interpolant to it. For example, in order to estimate derivatives or integrals of a function.

Applications? In mathematics, from number theory to information theory, and nearly every aspect of numerical analysis. Elsewhere, the methods are used in fields ranging from aircraft design to computer animation.

The main reference for this section is Chapter 6 of [SM03]. See also, Lectures 18–20 of [S96].

1.1.1 Polynomial Interpolation

Definition 1.1.1. $\mathcal{P}_n$ is the set of polynomials of degree less than or equal to n and real-valued coefficients, i.e., $p_n \in \mathcal{P}_n$ if

$$p_n(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

where $a_i \in \mathbb{R}$.

Examples:

Take notes:

It is particularly important to note that if p_n and q_n both belong to $\mathcal{P}_n$, then so too does their sum.

The Polynomial Interpolation Problem comes in two forms.

Polynomial Interpolation Problem 1 (PIP1): *Given is set of points* $x_0 < x_1 < \cdots < x_n$*, and a set of real numbers* $y_0, y_1, \ldots, y_n$*, find* $p_n \in \mathcal{P}_n$ *such that*

$$p_n(x_k) = y_k, \quad \textit{for } k = 0, 1, \ldots, n. \tag{1.1}$$

Polynomial Interpolation Problem 2 (PIP2): *Given is set of points* $x_0 < x_1 < \cdots < x_n$*, and a function* $f : [x_0, x_n] \to \mathbb{R}$*, find* $p_n \in \mathcal{P}_n$ *such that*

$$p_n(x_k) = f(x_k), \quad \textit{for } k = 0, 1, \ldots, n. \tag{1.2}$$

Clearly PIP2 is just PIP1 with $y_k = f(x_k)$.

The questions that we must ask (and answer) are

(i) Is there a solution to the polynomial interpolation problem.

(ii) Is it unique?

(iii) How do we find it?

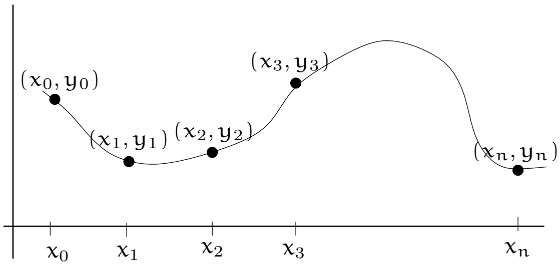


Fig. 1.2: A possible interpolation of the points in Figure 1.1

(iv) How accurate is it? That is, if f is the underlying function (i.e., $f(x_k) = y_k$), can we find an upper bound for

$$\max_{x_0 \leqslant x \leqslant x_n} \{|f(x) - p_n(x)|\}?$$

1.1.2 Exercises

Exercise 1.1. Suppose that $p \in \mathcal{P}_m$ and $q \in \mathcal{P}_n$.

(a) What is the maximum possible degree of $p + q$?

(b) What is the minimum possible degree of $p - q$?

(c) What is the maximum possible degree of pq?

Exercise 1.2. Find out what a *vector space* is. Convince yourself that $\mathcal{P}_n$ is a vector space.

Exercise 1.3. (a) Is it always possible to find a polynomial of degree 1 that interpolates the single point (x_0, y_0)? If so, how many such polynomials are there? Explain your answer.

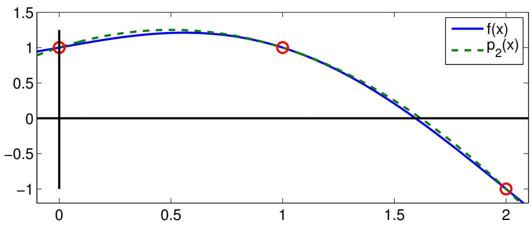
(b) Is it always possible to find a polynomial of degree 1 that interpolates the two points (x_0, y_0) and (x_1, y_1)? If so, how many such polynomials are there? Explain your answer.

(c) Is it ever possible to find a polynomial of degree 1 that interpolates the three points (x_0, y_0), (x_1, y_1), and (x_2, y_2)? If so, give an example.

1.2 Finding the polynomial

1.2.1 An example

Example 1.2.1. Show that the polynomial of degree 2 that interpolates $f(x) = 1 - x + \sin(\pi x/2)$ at the points $x_0 = 0$, $x_1 = 1$ and $x_2 = 2$ is $p_2 = -x^2 + x + 1$.



Take notes:

1.2.2 Uniqueness

It is not hard to convince ourselves that $-x^2 + x + 1$ is the solution to the PIP in Example 1.2.1. But how to we know we have found the *only* solution? More generally, *under what conditions is there exactly one polynomial that solves the PIP*?

Lemma 1.2.2. *If $p_n \in \mathcal{P}_n$ has $n+1$ zeros, then $p_n \equiv 0$ (i.e., $p_n(x) = 0$ for all x).*

Take notes:

Theorem 1.2.3 (There is a unique solution to the PIP)**.** *There is at most one polynomial of degree at most n that interpolates the $n+1$ points (x_0, y_0), (x_1, y_1), $\dots$, (x_n, y_n) where x_0, x_1, $\dots$, x_n are distinct.*

Take notes:

1.2.3 The Vandermonde matrix method

In Example 1.2.1 we gave a polynomial the solves a particular PIP, but this didn't tell us *how* to find the solution. It turns out that the most obvious approach may not be the best.

Suppose we are trying to solve the problem as follows: *find p_2 such that*

$$p_2(x_0) = y_0, \quad p_2(x_1) = y_1, \quad \textit{and} \quad p_2(x_2) = y_2.$$

Since $p_2(x)$ is of the form $a_0 + a_1 x + a_2 x^2$, this just amounts the finding the values of the coefficients a_0, a_1, and a_2. One might be tempted to solve for them using the system of equations

$$\begin{aligned} a_0 + a_1 x_0 + a_2 x_0^2 &= y_0 \\ a_0 + a_1 x_1 + a_2 x_1^2 &= y_1 \\ a_0 + a_1 x_2 + a_2 x_2^2 &= y_2. \end{aligned}$$

This is known as the *Vandermonde System*. [1]. Writing it in matrix-vector format we get:

$$\begin{pmatrix} 1 & x_0 & x_0^2 \\ 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ y_2 \end{pmatrix}, \quad \text{or} \quad V\mathbf{a} = \mathbf{y}. \tag{1.3}$$

But this may not be a good idea. (*You can skip the next bit if you did not take MA385*).

In MA385 we learned about the relationship between the *condition number* of a matrix, V, and the relative error in the (numerical) solution to a matrix-vector equation with V as the coefficient matrix. The condition number is

$$\kappa(V) = \|V\| \|V^{-1}\|,$$

for some subordinate matrix norm $\|\cdot\|$.

Example 1.2.4. ([S96, Lecture 18]) Suppose $x_0 = 100$, $x_1 = 101$ and $x_2 = 102$. Then it is not hard to check that

$$\|X\|_\infty = \max_i \sum_j |X_{ij}| = 10,507.$$

Also,

$$V^{-1} = \frac{1}{2} \begin{pmatrix} 10302 & -20400 & 10100 \\ -203 & 404 & -201 \\ 1 & -2 & 1 \end{pmatrix},$$

so $\|V^{-1}\|_\infty = 20401$. So $\kappa(V) = 214,353,307$.

[1] *Alexandre-Théophile Vandermonde* (France, 1735–1796) only began studying mathematics at the age of 35. Prior to that, he was a professional musician. Is credited for inventing the Vandermonde Matrix (and determinant), which he didn't do, but not for publishing the first paper in modern algebra, which he may have done.

1.2.4 Lagrange Interpolation

We'll now look at a much easier method for solving the Polynomial Interpolation Problem. As a by-product, we get a constructive proof of the existence of a solution to the PIP. (Here "constructive" means that we'll prove it exists by actually computing it).

Example 1.2.5. Consider the problem: *find* $p_3 \in \mathcal{P}_3$ *such that*

$$p_3(0) = 2, \quad p_3(1) = -1/2, \quad p_3(2) = 1, \quad p_3(3) = -1.$$

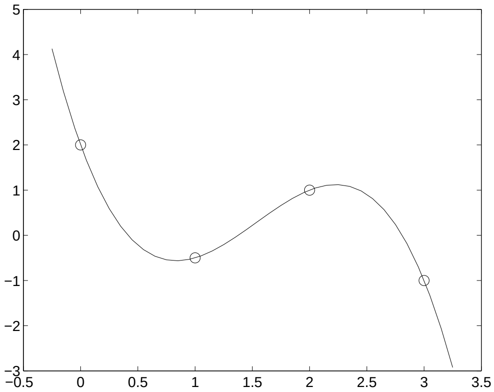


Fig. 1.1: A possible solution to Example 1.2.5

This is not so easy to solve; the following problem is easier: *find* $L_0 \in \mathcal{P}_3$ *such that*

$$L_0(0) = 1, \quad L_0(1) = 0, \quad L_0(2) = 0, \quad L_0(3) = 0.$$

Obviously, because L_0 is a cubic and has zeros at $x = 1, 2, 3$ it is of the form $L_0(x) = C(x-1)(x-2)(x-3)$. Then, choosing C so that $L_0(0) = 1$, we get

$$L_0(x) =$$

Similarly, let $L_1 \in \mathcal{P}_3$ be the cubic polynomial such that

$$L_1(0) = 0, \quad L_1(1) = 1, \quad L_1(2) = 0, \quad L_1(3) = 0,$$

then

$$L_1(x) =$$

In the same style, $L_2(x_i) = \begin{cases} 1 & i = 2 \\ 0 & i = 0, 1, 3 \end{cases}$ gives

$$L_2(x) =$$

Also, if we define

$$L_3(x_i) = \begin{cases} 1 & i = 3 \\ 0 & i = 0, 1, 2 \end{cases},$$

then clearly,

$$L_3(x) = \frac{(x-0)(x-1)(x-3)}{(3-0)(3-1)(3-2)} = \prod_{j=0, j\neq 3}^{n} \frac{(x-x_j)}{(x_3-x_j)}.$$

Plots of these polynomials are shown in Figure 1.2.

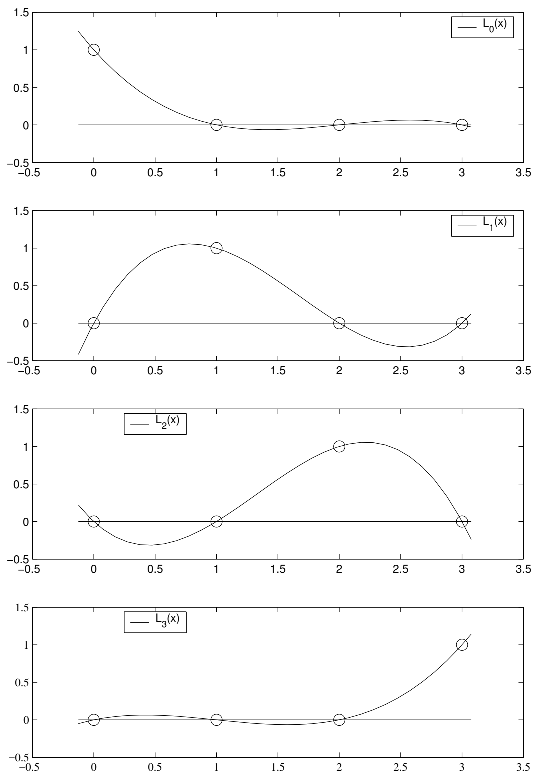


Fig. 1.2: The polynomials $L_0(x)$, $L_1(x)$, $L_2(x)$, and $L_3(x)$

Because each of L_0, L_1, L_2, and L_3 is a cubic, so too is any linear combination of them. So

$$p_3(x) = 2L_0(x) - (1/2)L_1(x) + (1)L_2(x) + (-1)L_3(x),$$

is a cubic. Furthermore

$$\begin{array}{rcllll}
p_3(0) & = & 2L_0(0) & -(1/2)L_1(0) & +(1)L_2(0) & +(-1)L_3(0) \\
& = & 2(1) & -(1/2)(0) & +(1)(0) & +(-1)(0) \\
& = & 2, & & & \\
p_3(1) & = & 2L_0(1) & -(1/2)L_1(1) & +(1)L_2(1) & +(-1)L_3(1) \\
& = & 2(0) & -(1/2)(1) & +(1)(0) & +(-1)(0) \\
& = & -1/2, & & & \\
p_3(2) & = & 2L_0(2) & -(1/2)L_1(2) & +(1)L_2(2) & +(-1)L_3(2) \\
& = & 2(0) & -(1/2)(0) & +(1)(1) & +(-1)(0) \\
& = & 1, & & & \\
p_3(3) & = & 2L_0(3) & -(1/2)L_1(3) & +(1)L_2(3) & +(-1)L_3(3) \\
& = & 2(0) & -(1/2)(0) & +(1)(0) & +(-1)(1) \\
& = & -1. & & &
\end{array}$$

Thus p_3 solves the problem in Example 1.2.5

1.2.5 The Lagrange Form

We can generalise this idea to solve any PIP using what is called *Lagrange*[2] interpolation.[3] We'll now look how to solve the general problem.

Definition 1.2.6. The **Lagrange Polynomials** associated with $x_0 < x_1 < \cdots < x_n$ is the set $\{L_i\}_{i=0}^n$ of polynomials in $\mathcal{P}_n$ such that

$$L_i(x_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}. \tag{1.4a}$$

and are given by the formula

$$L_i(x) = \prod_{j=0, j\neq i}^{n} \frac{x - x_j}{x_i - x_j}. \tag{1.4b}$$

Definition 1.2.7. The **Lagrange form of the Interpolating Polynomial**[a]

$$p_n(x) = \sum_{i=0}^{n} y_i L_i(x), \tag{1.5a}$$

or

$$p_n(x) = \sum_{i=0}^{n} f(x_i) L_i(x). \tag{1.5b}$$

[a]Take care not to confuse the *Lagrange Polynomials*, which are the L_i with the *Lagrange Interpolating Polynomial*, which is the p_n defined in (1.5).

Theorem 1.2.8 (Lagrange). *There exists a solution to the Polynomial Interpolation Problem and it is given by*

$$p_n(x) = \sum_{i=0}^{n} y_i L_i(x).$$

<u>`Take notes`</u>:

[2] 

Joseph-Louis Lagrange, born 1736 in Turin, died 1813 in Paris. He made great contributions to many areas of Mathematics, including *Calculus of Variations*, which we'll see a little of at the end of the semester. Image source: `http://jeff560.tripod.com/stamps.html`

[3]This formula was actually discovered by E. Waring in 1776; rediscovered by Euler in 1783, and published by Lagrange in 1795

Example 1.2.9. [SM03, E.g. 6.1] Write down the Lagrange form of the polynomial interpolant to the function $f(x) = e^x$ at interpolation points $\{-1, 0, 1\}$.

<u>`Take notes`</u>:

Figure 1.3 shows the solution to Example 1.2.9 (top) and the difference between the function e^x and its interpolant (bottom). It would be interesting to see how this error depends on

(i) the function (and its derivatives)

(ii) the number of points used.

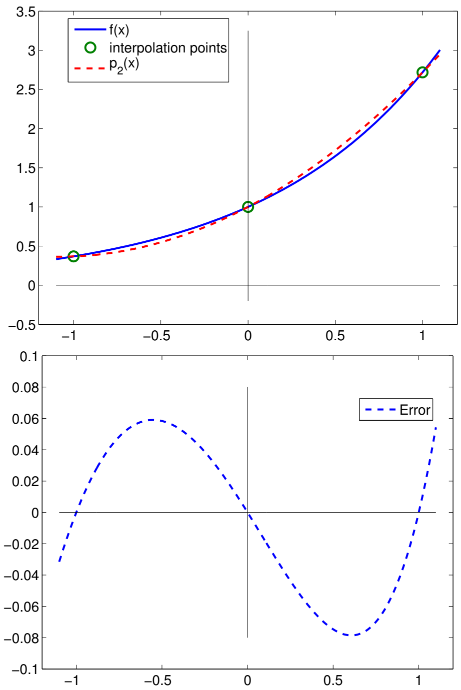


Fig. 1.3: The solution to Eg. 1.2.9 and the error

1.2.6 Some terminology

- The process of finding a polynomial, p_n, that solves the PIP is called *interpolation*.
- The solution, p_n, is called an *interpolant*.
- The difference between $f(x)$ and $p_n(x)$ is called the *interpolation error*.

1.2.7 Exercises

Exercise 1.4. Give a proof of Lemma 1.2.2 that is different from the one we did in class.

Exercise 1.5. The general form of the *Vandermonde* Matrix is

$$V_n = \begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{pmatrix}.$$

Its determinant is

$$\det(V_n) = \prod_{1 \leqslant i < j \leqslant n} (x_j - x_i). \tag{1.6}$$

Verify this for the 2×2 and 3×3 cases.

(Note that from formula (1.6) we can deduce directly that the PIP has a unique solution *if and only if* the points $x_0, x_1, \ldots, x_n$ are all distinct.)

Aside: Here is a hint for proving that in general

$$\det(V_n) = \prod_{1 \leqslant i < j \leqslant n} (x_j - x_i).$$

First note that $\det(V_n) = \det(V_n^T)$ and now consider the determinant of

$$V_n^T = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_0 & x_1 & x_2 & \cdots & x_n \\ x_0^2 & x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & & \vdots \\ x_0^n & x_1^n & x_2^n & \cdots & x_n^n \end{pmatrix}.$$

Using elementary row operations (which preserve the determinant), add to each row, the row above it multiplied by $-x_0$, to show that we need to compute the determinant of

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & x_1 - x_0 & x_2 - x_0 & \cdots & x_n - x_0 \\ 0 & x_1^2 - x_1 x_0 & x_2^2 - x_2 x_0 & \cdots & x_n^2 - x_n x_0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & x_1^n - x_1^{n-1} x_0 & x_2^n - x_2^{n-1} x_0 & \cdots & x_n^n - x_n^{n-1} x_0 \end{pmatrix}.$$

Now try to continue by induction...

Exercise 1.6. ⋆ (MA378 Semester 2 exam, 2015/2016) Let p_3 be the polynomial that interpolates $f(x) = \sin(x/2)$ at the points $x_0 = 0$, $x_1 = 1/3$, $x_2 = 2/3$ and $\mathbf{x_3 = 1}$. Use Cauchy's Theorem to give a bound for the error at $x = 1/2$. (You don't have to give a formula for p_3).[4].

Exercise 1.7. Find the polynomial p_1 that interpolates the function $f(x) = x^3$ at the points $x_0 = 0$ and $x_1 = a$. Find the point $\sigma \in [0, a]$ that maximises $|f(x) - p_1(x)|$, and hence compute $\max_{0 \leqslant x \leqslant a} |f(x) - p_1(x)|$.
Source: Chapter 6 of Süli and Mayers.

[4] An earlier version of this stated, incorrectly, that $x_4 = 1$

Exercise 1.8. ⋆ Show that

$$\sum_{i=0}^{n} L_i(x) = 1 \qquad \text{for all } x.$$

Exercise 1.9. Write down the Lagrange Form of p_2, the polynomial of degree 2 that interpolates the points $(0, 3)$, $(1, 2)$ and $(2, 4)$.
Source: Chapter 2 of Stoer and Bulirsch.

Exercise 1.10. Show that all the following represent the same polynomial (usually called the "Chebyshev Polynomial of Degree 3"), $T_3(x) = 4x^3 - 3x$.

(a) Horner form: $\big((4x + 0)x - 3\big)x + 0$.

(b) Lagrange form: $\displaystyle\sum_{k=0}^{3} \Big(\prod_{j=0, j\neq k}^{3} \frac{x - x_j}{x_k - x_j} \Big) (-1)^{k+1}$, where $x_0 = -1, x_1 = -1/2, x_2 = 1/2, x_3 = 1$.

(c) Recurrence relation: $T_0 = 1$, $T_1 = x$, and $T_n = 2xT_{n-1} - T_{n-2}$ for $n = 2, 3, \ldots$

(d) Trigonometric form: $T_3(x) = \cos\big(3\cos^{-1}(x)\big)$.

1.3 Polynomial Interpolation Errors

1.3.1 Introduction

In Example 1.2.9, we wrote down the polynomial of degree $n = 2$ interpolating $f(x) = e^x$ at $x_0 = -1$, $x_1 = 0$ and $x_2 = 1$.

We now want to investigate how, in general, error in polynomial interpolation depends on

(i) the function (and its derivatives)

(ii) the number of points used (or, equivalently, degree of the polynomial used).

The main ingredient we need is the following theorem.

Theorem 1.3.1 (Rolle's[5] Theorem). *Let* g *be a function that is continuous and differentiable on the interval* $[a, b]$. *If* $g(a) = g(b)$, *then there is at least one point* c *in* (a, b) *where* $g'(c) = 0$.

Our "proof" is by picture:[6]
Take notes:

1.3.2 Error estimate for n=0

The simplest case is when $n = 0$, so the interpolant is a constant, i.e., it is p_0 interpolating a function f at a point x_0. Here is one way we can deduce the *interpolation error*.[7]
Take notes:

It is important to understand what this formula is telling us:
Take notes:

1.3.3 Error estimates for $n \geqslant 1$

We will use this Rolle's Theorem to prove the most important theorem of NA2; it is used repeatedly through-out the course. It's often called the *Polynomial Interpolation Error Theorem*, but we'll call it *Cauchy's Theorem* for short.

First, we need to define an important polynomial.

Definition 1.3.2. The **Nodal Polynomial** π_{n+1} associated with the interpolation points that $a = x_0 < x_1 < \cdots < x_n = b$ is

$$\pi_{n+1}(x) = (x - x_0)(x - x_1)\ldots(x - x_n) = \prod_{i=0}^{n}(x - x_i).$$

Theorem 1.3.3 (Cauchy, 1840). *Suppose that* $n \geqslant 0$ *and* f *is a real-valued function that is continuous and defined on* $[a, b]$, *such that the derivative of* f *of order* $n + 1$ *exists and is continuous on* $[a, b]$. *Let* p_n *be the polynomial of degree* n *that interpolates* f *at the* $n + 1$ *points* $a = x_0 < x_1 < \cdots < x_n = b$. *Then, for any* $x \in [a, b]$ *there is a* $\tau \in (a, b)$ *such that*

$$f(x) - p_n(x) = \frac{f^{(n+1)}(\tau)}{(n+1)!}\pi_{n+1}(x). \tag{1.7}$$

Proof: Take notes:

[5] Michel Rolle, Born 1652 in Ambert, Basse-Auvergne (France), died 1719 in Paris. This is his most famous result.

[6] One can easily deduce Rolle's Theorem from the Mean Value Theorem (MVT). But since the standard proof of the MVT uses Rolle's Theorem, that would be cheating

[7] There is an easier way, involving the MVT, but the approach given here is easier to generalise

Example 1.3.4. Recall Theorem 1.2.9, where we wrote down the Lagrange form of the polynomial, p_2, that interpolates $f(x) = e^x$ at the points $\{-1, 0, 1\}$. Give a formula for $e^x - p_2(x)$.

Take notes:

Usually (and as in the above example), we can't calculate $f(x) - p_n(x)$ exactly from Formula (1.7), because we have no way of finding τ. However, we are typically not so interested in what the error is at some given point, but what is the maximum error over the whole interval $[x_0, x_n]$. That is given by:

Corollary 1.3.5. *Define*

$$M_{n+1} = \max_{x_0 \leqslant \sigma \leqslant x_n} |f^{(n+1)}(\sigma)|.$$

Then

$$|f(x) - p_n(x)| \leqslant \frac{M_{n+1}}{(n+1)!} |\pi_{n+1}(x)|. \tag{1.8}$$

Example 1.3.6. Let p_1 be the polynomial of degree 1 that interpolates a function f at distinct points x_0 and x_1. Letting $h = x_1 - x_0$, show that

$$\max_{x_0 \leqslant x \leqslant x_1} |f(x) - p_1(x)| \leqslant \frac{1}{8} h^2 M_2.$$

Take notes:

1.3.4 Exercise

Exercise 1.11. Let p_2 be the polynomial of degree 2 that interpolates a function f at the points x_0, x_1 and x_2. If $x_1 - x_0 = x_2 - x_1 = h$, show that

$$\max_{x_0 \leqslant x \leqslant x_2} |f(x) - p_2(x)| \leqslant \frac{1}{6} \frac{2}{3\sqrt{3}} h^3 M_3 = \frac{1}{9\sqrt{3}} h^3 M_3.$$

Hint: simplify the calculations by taking $t = x - x_1$, *writing* $(x - x_0)(x - x_1)(x - x_2)$ *in terms of* h *and* t.

1.4 Computing the polynomial interpolant

1.4.1 Synthetic Division

The Lagrange form of the polynomial interpolant has great theoretical importance. It is also useful for $n = 0, 1, 2$, say. It practice, for larger n, it is easy to write down the Lagrange form of the polynomial interpolant, but it is tedious (and computationally expensive) to work with. Evaluating each of the

$$L_i(x) = \prod_{j=0, j\neq i}^{n} \frac{x - x_j}{x_i - x_j}$$

at a given value of x requires $2n - 1$ multiplications. So therefore computing

$$p_n(x) = \sum_{i=0}^{n} y_i L_i(x) = \sum_{i=0}^{n} \Big(y_i \prod_{j=0, j\neq i}^{n} \frac{x - x_j}{x_i - x_j} \Big),$$

for a given value of x requires $2n^2 + 2n$ multiplications. On the other hand, suppose we have p_n in the standard form

$$p_n(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n.$$

At first glance it seems that this takes $(n^2 - n)/2$ multiplications, which is only a little better. However there is a faster way of doing this called *synthetic division*. As an example we'll consider how this works for the case $n = 3$. For the general case, please have a look at [S96, Lecture 19].

`Take notes`:

In summary,

- The Lagrange form is easy to find and difficult to evaluate.
- The standard form (i.e., $p_n = a_0 + a_1 x + \cdots + a_n x^N$), is hard to find, but easy to evaluate.
- So we want to find a form for the polynomial is easy to construct *and* efficient to evaluate.

There are plenty of possibilities – too many to mention – and we'll consider one of them: the *Newton Form*.

1.4.2 The Newton Form of the Interpolant

This section is just for your information: it will not be covered in class or be part of any assessment.

These notes are based on §2.1.3 of [SB92] and on [S96, Lecture 19].

Suppose we were to write the interpolating polynomial p_n as

$$\begin{aligned} p_n(x) = a_0 &+ a_1(x - x_0) \\ &+ a_2(x - x_0)(x - x_1) \\ &+ a_3(x - x_0)(x - x_1)(x - x_2) \\ &+ \cdots + a_n(x - x_0)(x - x_1)\cdots(x - x_{n-1}). \end{aligned}$$

Then it could actually be evaluated as

$$\begin{aligned} p_n(x) = a_0 + \\ (x - x_0)\Big(a_1 + (x - x_1)\big(a_2 + (x - x_2)(a_3 + \\ (\cdots + a_n(x - x_{n-1}))\cdots\big)\Big). \end{aligned} \tag{1.9}$$

This is called the *Newton Form of the Interpolating Polynomial.* [8]

Let π_k be the nodal polynomial $\pi_k = \prod_{i=0}^{k}(x - x_i)$. Then the Newton form is

$$p_n(x) = a_0 + a_1\pi_0 + a_2\pi_1 + a_3\pi_2 + \cdots + a_n\pi_{n-1}.$$

So how do we find the coefficients? Let

$$F[x_i] = f(x_i),$$

$$F[x_i, x_{i+1}] = \frac{F[x_{i+1}] - F[x_i]}{x_{i+1} - x_i}$$

$$\vdots$$

$$F[x_i, x_{i+1}, \dots, x_{i+k}] = \frac{F[x_{i+1}, x_{i+2}, \dots x_{i+k}] - F[x_i, x_{i+1}, \dots x_{i+k-1}]}{x_{i+k} - x_i}$$

Then

$$a_0 = f[x_0], \qquad a_1 = F[x_0, x_1], \quad \dots$$

$$a_n = F[x_0, x_1, \dots x_n].$$

Or, if you prefer:

$$p_n(x) = F[x_0] + \sum_{k=1}^{n} F[x_0, x_1, \dots, x_k]\pi_k(x). \tag{1.10}$$

[8] 
Sir Issac Newton Isaac Newton, 1643 - 1727, England. Easily one of the greatest scientist of all time. This image is of a Polish stamp commemorating the scientific achievements. Source: `http://jeff560.tripod.com/stamps.html`

It is not hard to show that the formula given in (1.10) yields the interpolating polynomial. The argument is inductive. Before we do that, we will demonstrate that it works for a particular example.

Example 1.4.1. Write down the Newton form of the polynomial p_2 that interpolates e^x at $x_0 = 0$, $x_1 = 1$ and $x_2 = 2$.

Solution: (For the Newton form)

	$F[x_i]$	$F[x_i x_{i+1}]$	$F[x_i x_{i+1} x_{i+2}]$
x_0	1		
		$e-1$	
x_1	e		$\frac{1}{2}(e^2-2e+1)$
		e^2-e	
x_2	e^2		

This gives

$$\begin{aligned} p_2(x) &= F[x_0] + (x-x_0)\big(F[x_0x_1] + (x-x_1)F[x_0x_1x_2]\big), \\ &= 1 + x\big((e-1) + (x-2)\frac{1}{2}(e^2-2e+1)\big). \end{aligned}$$

Written in the form of (1.10) above, this is

$$p_2(x) = 1 + (e-1)(x) + \frac{1}{2}(e^2-2e+1)(x)(x-1).$$

One can quickly check that this

1. is a polynomial of degree n,
2. interpolates e^x at $x=0$, $x=1$ and $x=2$.

It is not hard to show that the formula given in (1.10) yields the interpolating polynomial. The argument is inductive. Clearly its true for $n=0$, because in that case the interpolant is just the constant polynomial $p_0 = F[x_o] = f(x_0)$.

Next, suppose it is true for n points. Then let

$$\begin{aligned} p_n &\text{ interpolate } f(x) \text{ at } x = x_0, x_1, \dots x_n, \\ q_{n-1}(x) &\text{ interpolate } f(x) \text{ at } x = x_0, x_1, \dots x_{n-1}, \\ r_{n-1}(x) &\text{ interpolate } f(x) \text{ at } x = x_1, x_2 \dots x_n. \end{aligned}$$

We must convince ourselves that

$$p_n(x) = q_{n-1}(x) + \frac{x-x_0}{x_n-x_0}\big(r_{n-1}(x) - q_{n-1}(x)\big),$$

First, since both q_{n-1} and r_{n-1} are polynomials of degree $n-1$, it follows that the expression on the right is indeed a polynomial of degree n.

Next, check that it is the interpolant of f at x_0, x_1, ..., x_n. For the first point, x_0:

$$\begin{aligned} p_n(x_0) &= q_{n-1}(x_0) + \frac{x_0-x_0}{x_n-x_0}\big(r_{n-1}(x_0) - q_{n-1}(x_0)\big) \\ &= q_{n-1}(x_0) = f(x_0). \end{aligned}$$

For the last point, x_n:

$$\begin{aligned} p_n(x_n) &= q_{n-1}(x_n) + \frac{x_n-x_0}{x_n-x_0}\big(r_{n-1}(x_n) - q_{n-1}(x_n)\big) \\ &= q_{n-1}(x_n) + \big(r_{n-1}(x_n) - q_{n-1}(x_n)\big) \\ &= r_{n-1}(x_n) = f(x_n). \end{aligned}$$

And for every other point, x_i, $i = 1, \dots, n-1$,

$$\begin{aligned} p_n(x_i) &= q_{n-1}(x_i) + \frac{x_i-x_0}{x_n-x_0}\big(r_{n-1}(x_i) - q_{n-1}(x_i)\big) \\ &= f(x_i) + \frac{x_i-x_0}{x_n-x_0}\big(f(x_i) - f(x_i)\big) = f(x_i). \end{aligned}$$

Finally, we need to compare the coefficient of x^n on the left- and right-hand sides of this equations.

The coefficient of x^{n-1} in q_{n-1} is $F[x_0, x_1, \dots, x_{n-1}]$. The coefficient of x^{n-1} in r_{n-1} is $F[x_1, x_2, \dots, x_n]$. So the coefficient of x^n in p_n is

$$\frac{1}{x_n-x_0}\big(F[x_1, x_2, \dots, x_n] - F[x_0, x_1, \dots, x_{n-1}]\big),$$

as required.

1.4.3 Exercise

Exercise 1.12. Do Exercise 6.3 from Süli and Mayers, *An Introduction to Numerical Analysis*.

1.5 Hermite Interpolation

1.5.1 The idea

Hermite[9] interpolation is a variant on the standard Polynomial Interpolation Problem: we seek a polynomial that not only agrees with a given function f at the interpolation points, but its first derivative also matches f' at those points. We are not that interested in this problem for its own sake, but the idea recurs again in the sections in piecewise polynomial interpolation and Gaussian quadrature.

Formally, the problem is

The Hermite Polynomial Interpolation Problem (HPIP)

Given a set of interpolation points $x_0 < x_1 < \cdots < x_n$ *and a continuous, differentiable function* f*, find* $p_{2n+1} \in \mathcal{P}_{2n+1}$ *such that*

$$p_{2n+1}(x_i) = f(x_i) \quad \textit{and} \quad p'_{2n+1}(x_i) = f'(x_i).$$

It is not hard to see that if there is a solution to this problem, then it is unique:

Take notes:

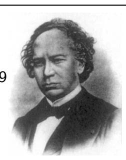
1.5.2 Constructing Hermite Interpolants

It is possible to solve this problem using an extension of the Lagrange Polynomial approach of §1.2.4. Given the Lagrange Polynomials L_i as defined in (1.4) let

$$H_i(x) = [L_i(x)]^2(1 - 2L_i'(x_i)(x - x_i)),$$

$$K_i(x) = [L_i(x)]^2(x - x_i).$$

Examples of two such functions are shown in Figure 1.1.

9  Charles Hermite, France, 1822–1901. A part from this form of interpolation, his contributions to Mathematics included the first proof that e is transcendental. His methods were later used to show that π is transcendental.

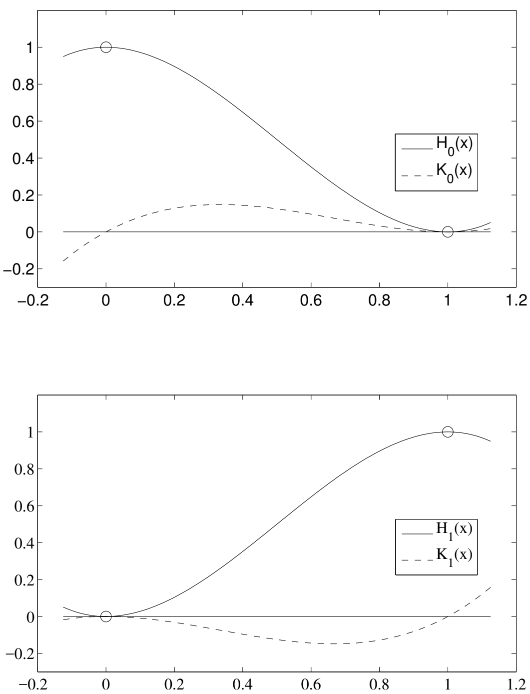


Fig. 1.1: Hermite bases functions H_0, K_0 (top) and H_1, K_1 (bottom) for $n = 1$, $x_0 = 0$ and $x_1 = 1$

We can show that, for $i, k = 0, 1, \ldots n$,

$$H_i(x_k) = \begin{cases} 1 & i = k \\ 0 & i \neq k \end{cases} \qquad H_i'(x_k) = 0 \ \forall k,$$

$$K_i(x_k) = 0, \qquad K_i'(x_k) = \begin{cases} 1 & i = k \\ 0 & i \neq k \end{cases}$$

We'll show this for H and H', and leave the corresponding results for K and K' to Exercise 1.16.

Take notes:

Armed with these identities, we can now show that the solution to the HPIP exists and is given by

$$p_{2n+1}(x) = \sum_{i=0}^{n} \big(f(x_i)H_i(x) + f'(x_i)K_i(x)\big).$$

(Again, see Exercise 1.16 for details).

Example 1.5.1. Find the polynomial of degree 3 that interpolates $\exp(x^2)$, and its first derivative, at $x_0 = 0$ and $x_1 = 1$. (See Figure 1.2).

Take notes:

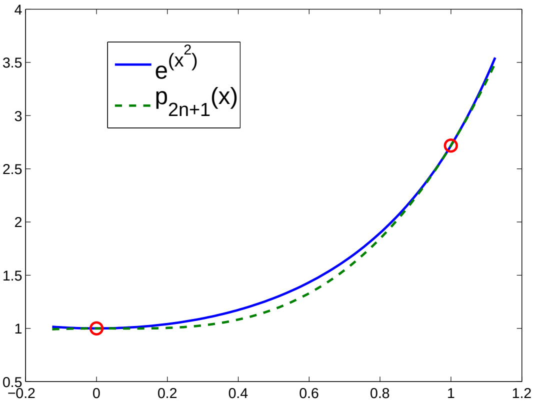


Fig. 1.2: The Hermite interpolant to $e^{(x^2)}$ at $x = 0$ and $x = 1$

Theorem 1.5.2. *Let be a real-valued function that is continuous and defined on* $[a, b]$*, such that the derivatives of* f *of order* $2n + 2$ *exist and are continuous on* $[a, b]$*. Let* p_{2n+1} *be the Hermite interpolant to* f*. Then, for any* $x \in [a, b]$ *there is an* $\tau \in (a, b)$ *such that*

$$f(x) - p_{2n+1}(x) = \frac{f^{(2n+2)}(\tau)}{(2n+2)!}[\pi_{n+1}(x)]^2.$$

We won't do a proof of this in class. However, later in this course we'll be interested in the particular example of finding p_3 the cubic Hermite Polynomial Interpolant to a function f at the points x_0 and x_1. Also, see Exercise 1.13.

1.5.3 Exercises

Exercise 1.13. $\star$ For *just* the case $n = 1$, state and prove an appropriate version of Theorem 1.5.2 (i.e., error in the Hermite interpolant). Use this to find a bound for $\|f - p_3\|_{[x_0,x_1]}$ in terms of f and $h = x_1 - x_0$. (Here $\|g\|_{[x_0,x_1]}$ is short-hand for $\max_{x_0 \leqslant x \leqslant x_1} |g(x)|$.)

Exercise 1.14. Let $n = 2$ and $x_0 = -1$, $x_0 = 1$ and $x_1 = 1$. Write out the formulae for H_i and K_i for $i = 0, 1, 2$ and give a rough sketch of each of these six functions that shows the value of the function and its derivative at the three interpolation points.

Exercise 1.15. Do Exercise 6.6 from from Süli and Mayers, *An Introduction to Numerical Analysis*.

Exercise 1.16. $\star$ Let L_0, L_1, ..., L_n be the usual Lagrange polynomials for the set of interpolation points $\{x_0, x_1, \dots, x_n\}$. Now define

$$H_i(x) = [L_i(x)]^2(1 - 2L_i'(x_i)(x - x_i)),$$

and

$$K_i(x) = [L_i(x)]^2(x - x_i).$$

We saw in class that, for $i, k = 0, 1, \dots n$,

$$H_i(x_k) = \begin{cases} 1 & i = k \\ 0 & i \neq k \end{cases} \qquad H_i'(x_k) = 0.$$

Show that, for $i, k = 0, 1, \dots n$,

$$K_i(x_k) = 0, \qquad K_i'(x_k) = \begin{cases} 1 & i = k \\ 0 & i \neq k \end{cases}$$

Conclude that the solution to the Hermite Polynomial Interpolation Problem is

$$p_{2n+1}(x) = \sum_{i=0}^{n} \big(f(x_i)H_i(x) + f'(x_i)K_i(x)\big).$$

Exercise 1.17. Write down that formula for q_3, the *Hermite* polynomial that interpolates $f(x) = \sin(x/2)$, and its derivative, at the points $x_0 = 0$ and $x_1 = 1$. Give an upper bound for $|f(1/2) - q_3(1/2)|$. How does this compare with the bound in Exercise 1.6?

Exercise 1.18. (This exercise is based on Exer 6.5 from Süli and Mayers' *Introduction to Numerical Analysis*). Consider the following problem.

> Take $n+1$ distinct interpolation points $x_0 < x_1 < \cdots < x_n$. Let p_{2n+1} be the polynomial of degree $2n + 1$ with the property that
>
> $$p_{2n+1}(x_i) = f(x_i),$$
>
> and
>
> $$p''_{2n+1}(x_i) = f''(x_i).$$

In general this problem does *not* have a unique problem.

(i) Explain briefly but carefully why the arguments, based on Rolle's Theorem, used to prove **uniqueness** of solutions to the HPIP, will not work here.

(ii) Show that there is no $p_5(x)$ that solves this problem when

- $x_0 = -1$, $x_1 = 0$, $x_2 = 1$.
- $f(-1) = 1$, $f(0) = 0$, $f(1) = 1$.
- $f''(-1) = 0$, $f''(0) = 0$, $f''(1) = 0$.

1.6 Wrap-up

1.6.1 Convergence & Runge's Example

The celebrated Weierstrass approximation theorem states that, given f and a positive number ε, there is a polynomial p such that

$$\max_{x\in[a,b]} |f(x)-p(x)| := \|f-p\|_\infty \leqslant \varepsilon.$$

Now suppose that f is a continuous function on $[a,b]$ and that $\{p_n\}_{n=0}^{\infty}$ is a sequence of polynomials that interpolate f at $n+1$ **equally spaced** points. One might be inclined to believe that

$$\lim_{n\to\infty} \|f-p_n\|_\infty = 0.$$

Equivalently, recalling (1.8):

$$|f(x)-p_n(x)| \leqslant \frac{M_{n+1}}{(n+1)!}|\pi_{n+1}(x)|,$$

one might expect that

$$\lim_{n\to\infty} \max_{x\in[a,b]} \frac{M_{n+1}}{(n+1)!}|\pi_{n+1}(x)| = 0.$$

In order words, we might think that, in order to find an interpolating polynomial that is as accurate as we would like, we just need to choose large enough n.

And some times it might work. For example, suppose that $a=-5$, $b=5$, and $f(x)=e^{\sin(x/2)}$. In Table 1.1 the errors for successive interpolants are shown. A few of these are shown in Figure 1.3.

Table 1.1: Errors in polynomial interpolants to $e^{\sin(x/2)}$ on $[-5,5]$

n	$\|f-p_n\|_\infty$
2	1.27e-00
4	2.94e-01
6	8.39e-02
8	5.75e-02
16	1.07e-03

However, there is a famous example of a simple function that cannot be successfully interpolated in this manner. This is *Runge's Example*:

$$f(x) = \frac{1}{1+x^2} \qquad \text{on } [-5,5].$$

The errors are shown in Table 1.2. Notice that they *increase* with n. We don't have the scope to go into *why* this is the case (but see the notes from class on a heuristic explanation).

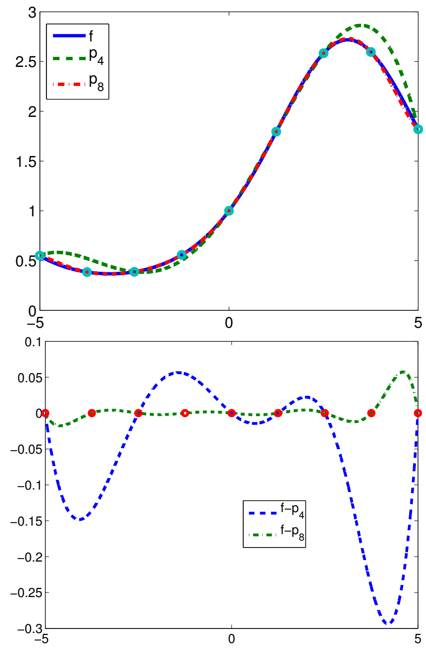


Fig. 1.3: Polynomial interpolants to $e^{\sin(x/2)}$ on $[-5,5]$, and their errors (bottom)

Table 1.2: Errors in interpolants to $1/(1+x^2)$ on $[-5,5]$

n	$\|f-p_n\|$
2	0.65
4	0.44
6	0.62
8	1.05
16	14.39
20	59.66
22	122.91
24	257.21

1.7 Where to from here?

So now it looks like polynomial interpolation is bad, at least on equidistant points. However, our experience in Lab 1 might lead us to be more optimistic: we were able to find a set of points that made the approximation as good we wanted (until round-off error dominated).

Unfortunately, just because we have a good set of points for interpolating one particular function, it does not follow that that set is good for every continuous function: this is *Faber's Theorem*[10]. This has often led numerical analysts to abandon the idea of interpolation by high-order polynomials completely.

However, there is a set of points that are useful, if f is smooth enough: the Chebyshev points of Lab 1. If you are interested, there please look at the essay *Inverse Yogiisms* by Lloyd N. (Nick) Trefethen, Notices AMS, Dec. 2016.[11] To investigate this numerically in MATLAB, try

[10] Georg Faber (18771966)

[11] `http://people.maths.ox.ac.uk/trefethen/essays.html` or `http://www.ams.org/publications/journals/notices/`

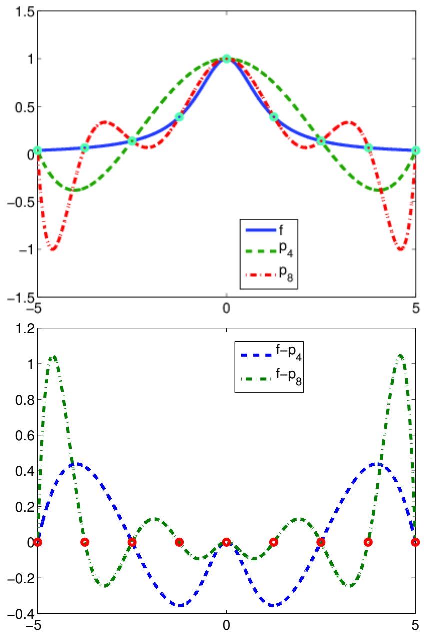


Fig. 1.4: Polynomial interpolants to $1/(1+x^2)$ on $[-5, 5]$

exploring the Chebfun toolbox.

The approach we will take is different. In (1.3.6) we saw that

$$\max_{x_0 \leqslant x \leqslant x_1} |f(x) - p_1(x)| \leqslant \frac{1}{8} h^2 M_2.$$

So, assuming M_2 is bounded (which is reasonable), we can make p_1 as close to f as we would like by taking a small enough interval $[x_0, x_1]$. The next section of this module is devoted to seeing how this can be used in theory and practice.

201611/rnoti-p1281.pdf