Notes on Polynomial Interpolation

2D1250, Tillämpade numeriska metoder II

Olof Runborg

March 19, 2003

These notes expand the material on polynomial interpolation in Heath, filling in gaps with further explanation and proofs of some statements.

1 Polynomial Interpolation

The interpolation problem that we are interested in is the following: Given a continuous function $f(x)$ defined on the interval $[a, b]$ and n distinct *nodes* (x-coordinates) satisfying

$$a \leq x_1 < x_2 < \cdots < x_n \leq b,$$

find an interpolation polynomial $p(x)$ such that

$$f(x_j) = p(x_j), \qquad j = 1, \ldots, n.$$

This is sometimes called *Lagrange interpolation* to distinguish it from Hermite interpolation, see below. For the basics, see Chapter 7 in Heath.

As in Heath, we let $\mathbb{P}_n$ be the (function) space of polynomials of degree at most n. We let $C([a, b])$ denote the set of continuous functions on the interval $[a, b]$. Similarly, we let $C^n([a, b])$ be the n times continuously differentiable functions on $[a, b]$.

1.1 Existence and uniqueness

The main existence result says that the interpolation polynomial above always exists and that there is only one polynomial of degree (at most) $n-1$ that interpolates n given, distinct, nodes. To be precise, we write this as a theorem.

Theorem 1 *For any set of n distinct nodes $x_1 < \cdots < x_n$ and associated function values $f(x_j)$, there exists a unique polynomial $p(x)$ in $\mathbb{P}_{n-1}$ for which*

$$f(x_j) = p(x_j), \qquad j = 1, \ldots, n. \tag{1}$$

Proof: There is a proof in the beginning of Section 7.3.1 in Heath (not available in the old edition). Another proof is as follows. Introduce the *Lagrange polynomials* (aka fundamental or characteristic polynomials) for the nodes,

$$\ell_j(x) := \prod_{\substack{i=1 \\ i \neq j}}^{n} \frac{x - x_i}{x_j - x_i}, \qquad j = 1, \ldots, n. \tag{2}$$

We claim that the interpolation polynomial is given by

$$p(x) = \sum_{i=1}^{n} f(x_i)\ell_i(x). \tag{3}$$

First, since there are $n-1$ factors in the product defining the Lagrange polynomials in (2), each of them belong to $\mathbb{P}_{n-1}$. A sum of polynomials in $\mathbb{P}_{n-1}$ is still in $\mathbb{P}_{n-1}$, so $p(x) \in \mathbb{P}_{n-1}$. Second, it is easy to see that the Lagrange polynomials have the property

$$\ell_i(x_j) = \delta_{i-j} := \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases}$$

Therefore, the polynomial satsifies (1):

$$p(x_j) = \sum_{i=1}^{n} f(x_i)\ell_i(x_j) = \sum_{i=1}^{n} f(x_i)\delta_{i-j} = f(x_j).$$

Finally, to show the uniqueness, suppose there is another polynomial $\bar{p}(x) \in \mathbb{P}_{n-1}$ that also satisfies (1). Then $q(x) := p(x) - \bar{p}(x)$ belongs to $\mathbb{P}_{n-1}$ and $q(x_j) = p(x_j) - \bar{p}(x_j) = f(x_j) - f(x_j) = 0$ for all j. But a polynomial in $\mathbb{P}_{n-1}$ with n distinct zeros must be identically zero, hence $p(x) \equiv \bar{p}(x)$. $\square$

1.2 Interpolation error

A theoretically useful expression for the pointwise interpolation error is contained in the following theorem.

Theorem 2 *Suppose $f \in C^n([a, b])$. Let $p(x)$ be the unique polynomial in $\mathbb{P}_{n-1}$ interpolating $f(x)$ at the nodes $\{x_j\}$ with $a = x_1 < x_2 < \cdots < x_n = b$. Then, for each $x \in [a, b]$ there is a $\xi \in [a, b]$ (depending on x), such that*

$$f(x) - p(x) = \frac{f^{(n)}(\xi)}{n!}(x - x_1)(x - x_2)\cdots(x - x_n). \tag{4}$$

Moreover, if $h = \max_j x_{j+1} - x_j$,

$$||f - p||_\infty := \sup_{a \leq x \leq b} |f(x) - p(x)| \leq \frac{h^n}{4n}||f^{(n)}||_\infty. \tag{5}$$

Proof: We start by showing (4). It is trivially true when $x = x_j$ for some j ("$0 = 0$"). We therefore fix an x such that $x \neq x_j$ for all j, and set

$$g(t) := f(t) - p(t) - \omega(t)\frac{f(x) - p(x)}{\omega(x)}$$

where $\omega(t) = (t - x_1)\cdots(t - x_n)$. We note that $g(t = x) = 0$ and $g(t = x_j) = 0$ for all j, so $g(t)$ has at least $n + 1$ zeros. By Rolle's theorem, applied to successive zero crossings, $g'(t)$ therefore has at least n zeros. Since $g \in C^n$ (because $f \in C^n$) we can use this argument inductively n times and conclude that $g^{(n)}(t)$ has at least one zero in $[a, b]$, which we denote ξ. Hence,

$$0 = g^{(n)}(\xi) = f^{(n)}(\xi) - p^{(n)}(\xi) - \omega^{(n)}(\xi)\frac{f(x) - p(x)}{\omega(x)}.$$

But $p^{(n)}(x) \equiv 0$ since $p \in \mathbb{P}_{n-1}$ and $\omega^{(n)}(x) \equiv n!$ since it has leading coefficient one and belongs to $\mathbb{P}_n$. We obtain

$$f^{(n)}(\xi) - n!\frac{f(x)-p(x)}{\omega(x)} = 0,$$

from which (4) follows. Finally, suppose $x_j \leq x < x_{j+1}$ for some j with $1 \leq j \leq n-1$. Then

$$\begin{aligned}
|\omega(x)| &= |x-x_1|\cdots|x-x_j|\cdots|x-x_n| \\
&\leq jh\,(j-1)h\cdots 2h\,(x-x_j)(x_{j+1}-x)\,2h\cdots(n-j)h \\
&= (x-x_j)(x_{j+1}-x)h^{n-2}j!(n-j)! \\
&\leq \frac{(x_{j+1}-x_j)^2}{4}h^{n-2}j!(n-j)! \leq \frac{h^n}{4}(n-1)!
\end{aligned}$$

This estimate, together with (4), gives (5). $\square$
Some remarks:

- The estimates (4) and (5) do not necessarily mean that the error $||f-p||_\infty$ decreases to zero when we interpolate with more and more points (increase n). Eventhough h may go to zero the derivatives of f, i.e. $||f^{(n)}||_\infty$, can grow rapidly and prevent convergence. Equidistant interpolation of *Runge's function* is a striking example showing that this can indeed happen. (See Fig. 7.7 in Heath.) In fact, high order polynomial interpolation on equally spaced nodes is notoriously bad, in particular close to the end points of the interval.

- When the interpolated function f is *analytic* around the interval, with poles sufficiently far away, the error *does* decrease to zero with n for equidistant interpolation. This is because $||f^n||_\infty$ cannot grow too fast with n for analytic functions. Let us make a simple derivation. Suppose the interval is $[-a,a]$ with $a+\varepsilon < R$ for some $\varepsilon > 0$ and R is the distance from the origin to the closest pole (the radius of convergence). A defining property of analytic functions is that

$$\sup_{|z|\leq a} |f^{(n)}(z)| \leq C_\varepsilon n! r^{-n}, \qquad r = R - a - \varepsilon,$$

where C_ε is a constant that depends on ε but not on n. Since $h = 2a/(n-1)$ for equidistant interpolation on n nodes, Theorem 2 and Stirling's formula (i.e. $(n-1)! \leq n^n e^{-(n-1)}$) yield

$$\sup_{a\leq x\leq b} |f(x)-p(x)| \leq C_\varepsilon \left(\frac{2a}{n-1}\right)^n \frac{1}{4n} n! r^{-n} \leq \frac{C_\varepsilon e}{4}\left(\frac{n}{n-1}\right)^n \left(\frac{2a}{re}\right)^n.$$

Noting that $1+x \leq \exp(x)$ we have for $n > 1$

$$\sup_{a\leq x\leq b} |f(x)-p(x)| \leq \frac{C_\varepsilon e}{4}\left[\exp\left(\frac{1}{n-1}\right)\right]^n \left(\frac{2a}{re}\right)^n \leq \frac{C_\varepsilon e^3}{4}\left(\frac{2a}{re}\right)^n.$$

The error hence tends to zero with n provided $2a < re$, which is equivalent to

$$R > \left(1+\frac{2}{e}\right)a + \varepsilon \approx 1.74\,a + \varepsilon.$$

This is not a sharp bound, and equidistant polynomial interpolation converges also for smaller R. The important point, however, is that we always get convergence when R is sufficiently large. For Runge's example, $a = 1$ but R is too small: since $\pm 0.2i$ are the poles, $R = 0.2$. On the other hand, functions which are analytic everywhere, *entire* functions such as $\exp(z)$, have $R = \infty$ and they are particulary well approximated by polynomials, c.f. exercise 1, lab 2.

- Discouraged by Runge's example, one may wonder if it is at all possible to find polynomials such that $||f - p_n||_\infty \to 0$ when f is not analytic. The answer is, however, positive. It is given by the Weierstrass approximation theorem:

  Theorem 3 *For any function $f \in C([a,b])$ there is a sequence of polynomials $p_0, p_1, \ldots,$ such that*
  $$\lim_{n\to\infty} ||f - p_n||_\infty = 0.$$

  Hence, even when f is merely continuous it can always be approximated arbitrarily well by a polynomial. One can also show that for each continuous function $f \in C([a,b])$ and degree n, there is a unique *polynomial of best approximation* $p_n^* \in \mathbb{P}_n$ such that
  $$||f - p_n^*||_\infty = \min_{p\in\mathbb{P}_n} ||f - p||_\infty.$$
  and that p_n^* interpolates the function at the $n+1$ *best nodes* $a \le x_1^* < \cdots < x_{n+1}^* \le b$,
  $$p^*(x_j^*) = f(x_j^*), \quad j = 1, \ldots, n+1.$$
  The best polynomial is thus interpolating, but unfortunately the location of the best nodes depends on f which makes them inconvenient to use.

2 Hermite interpolation

In Hermite interpolation also the *derivatives* of the function are interpolated. The interpolation problem can be stated as follows. Given a function $f \in C^M([a,b])$ and n distinct *nodes*,
$$a \le x_1 < x_2 < \cdots < x_n \le b,$$
find an interpolation polynomial $p(x)$ such that
$$f^{(k)}(x_j) = p^{(k)}(x_j), \qquad j = 1, \ldots, n, \qquad k = 0, \ldots, m_j,$$
where $m_j \le M$. The simplest form is so called *osculatory* interpolation where $m_j = 1$ for all j,
$$f(x_j) = p(x_j), \qquad f'(x_j) = p'(x_j), \qquad j = 1, \ldots, n.$$
The previous results for Lagrange interpolation generalize to Hermite interpolation. We have

Theorem 4 (Existence and uniqueness) *If the nodes are distinct, then there exists a unique polynomial $p(x)$ in $\mathbb{P}_{N-1}$, with $N = \sum_{j=1}^n (m_j + 1)$, that solves the Hermite interpolation problem above.*

and

Theorem 5 (Interpolation error) *Let $p(x)$ be the Hermite interpolation polynomial. Under the assumptions above, and if $f \in C^N([a,b])$ with $N = \sum_{j=1}^n (m_j+1)$, there is a $\xi \in [a,b]$ (depending on x), such that*
$$f(x) - p(x) = \frac{f^{(N)}(\xi)}{N!}(x - x_1)^{(m_1+1)}(x - x_2)^{(m_2+1)} \cdots (x - x_n)^{(m_n+1)}.$$

3 Alternatives to High Order Equidistant Interpolation

As mentioned above, high order polynomial interpolation on equally spaced nodes is notoriously bad, in particular close to the end points of the interval. Two ways of improving the interpolation are to use low order piecewise polynomial interpolation, or to use nodes that are not equidistant.

3.1 Piecewise polynomial interpolation

Here we divide the interval $I = [a, b]$ into a set of m subintervals separated by the $m + 1$ *knots* $\{x_j\}$,

$$I_j = [x_j, x_{j+1}], \qquad a = x_1 < x_2 < \cdots < x_{m+1} = b.$$

We then use polynomial interpolation of order k in each of the subintervals, usually with $k \ll m$. Let $q_j(x) \in \mathbb{P}_k$ be the polynomial in interval I_j and let the interpolation nodes in I_j be equidistant. We then have

$$f(x_{i,j}) = q_j(x_{i,j}), \qquad x_{i,j} = x_j + (i-1)\delta, \qquad \delta = \frac{x_{j+1} - x_j}{k}, \qquad i = 1, \dots k+1.$$

Finally, we define $\tilde{p}_{k,m}(x) \in C([a, b])$ to be the piecewise polynomial function defined by

$$\tilde{p}_{k,m}(x) = q_j(x), \qquad x \in I_j.$$

The total number of nodes on which we interpolate is now $n = km + 1$. In fact, calling $y_{(j-1)k+i} = x_{i,j}$ we have have that

$$f(y_\ell) = \tilde{p}_{k,m}(y_\ell), \qquad a = y_1 < y_2 < \cdots < y_n = b.$$

Let us now assume that the subintervals are all of the same size, $x_{j+1} - x_j = (b - a)/m$. Then the nodes $\{y_\ell\}$ are equidistant and $y_\ell = (\ell - 1)\delta$ with $\delta = (b - a)/(n - 1)$. It is then easy to see that the uniform interpolation error goes to zero with n if k is fixed. Suppose that $f \in C^{k+1}([a, b])$. Then formula (5) in Theorem 2 gives

$$||f - \tilde{p}_{k,m}||_\infty = \max_{1 \le j \le m} ||f - q_j||_\infty \le \frac{\delta^{k+1} ||f^{(k+1)}||_\infty}{4(k+1)} \le \frac{(b-a)^{k+1} ||f^{(k+1)}||_\infty}{4(k+1)(n-1)^{k+1}} \le C_k n^{-(k+1)}, \tag{6}$$

for some constant C_k independent of n. We say that the *order of accuracy* (in function evaluations n) is $k + 1$.

Example: Piecewise linear interpolation. This is when the interpolation points are simply connected by straight lines. Then $k = 1$ and the error is $\mathcal{O}(n^{-2})$. Similarly, for quadratic interpolation the error is $\mathcal{O}(n^{-3})$.

Note that, in this case when we interpolate on equidistant nodes there is no problem with convergence. This should be contrasted to the case when we used high order equidistant interpolation, where the Runge phenomenon could prevent convergence. With piecewise interpolation the error always decays uniformly to zero when we take more points, even for low order interpolation, i.e. small k. This is important since interpolation on equidistant nodes is of substantial practical interest. Function data is typically obtained in this way, e.g. from measurements or from computations that use regular grids.

3.1.1 Piecewise Hermite interpolation

The piecewise interpolant $\tilde{p}_{k,m}$ introduced above is continuous at the subinterval boundaries (at the knots x_j), but even for large k its derivatives are typically discontinuous. It is thus not smooth which may sometimes be a problem. One solution to this is to use piecewise Hermite interpolation, where we let $q_j(x)$ in the description above satisfy

$$f^{(\ell)}(x_j) = q_j^{(\ell)}(x_j), \qquad f^{(\ell)}(x_{j+1}) = q_j^{(\ell)}(x_{j+1}), \qquad j = 1, \ldots, m, \qquad \ell = 0, \ldots, k.$$

Hence, in each subinterval we pick the polynomial that agrees with the function f and its first m derivatives at the subinterval boundaries. Then by Theorem 4 we must take $q_j \in \mathbb{P}_{N-1}$ with $N = 2k+2$. We still let $\tilde{p}_{k,m}$ denote the piecewise polynomial (eventhough k now signifies something else). By the conditions above it is clear that $\tilde{p}_{k,m} \in C^k([a,b])$ and it is hence more smooth than for Lagrange interpolation. In this case we also get error estimates for the derivatives of $\tilde{p}_{k,m}$,

$$||f^{(\ell)} - \tilde{p}_{k,m}^{(\ell)}||_\infty \leq C_{\ell k} m^{-(N-\ell)} ||f^{(N)}||_\infty, \qquad \ell = 0, \ldots, k. \tag{7}$$

provided $f \in C^N([a,b])$. (Here $C_{\ell k}$ is a constant that depends on ℓ, k but not on m.) The order of accuracy for $\tilde{p}_{k,m}$ is thus N and for each derivative the order decreases by one.

3.1.2 Cubic splines

Another possibility to obtain smoother interpolants, is to use *splines* which are piecewise polynomials with continuity requirements imposed on their derivatives at the knots. The most common example is the *cubic spline.* Here we let $q_j \in \mathbb{P}_3$ satisfy

$$f(x_j) = q_j(x_j), \qquad f(x_{j+1}) = q_j(x_{j+1}), \qquad j = 1, \ldots, m,$$

and

$$q_j'(x_j) = q_{j-1}'(x_j), \qquad q_j''(x_j) = q_{j-1}''(x_j), \qquad j = 2, \ldots, m,$$

and some boundary conditions at $x_1 = a$ and $x_{m+1} = b$, for instance $q_1''(a) = q_m''(b) = 0$. We let $s_m(x)$ denote the full spline function,

$$s_m(x) = q_j(x), \qquad x \in I_j.$$

Note that we do not need to know the derivatives of f at the knots in this case. The extra conditions compared to Lagrange interpolation, are instead used to make both the first and and second derivative continuous everywhere, so that $s_m \in C^2([a,b])$. In order to determine q_j we must solve a tridiagonal linear system of equations at a cost of $\mathcal{O}(m)$ operations. Also for splines we get an error estimate of the form (7),

$$||f^{(\ell)} - s_m^{(\ell)}||_\infty \leq C_\ell m^{-(4-\ell)} ||f^{(4)}||_\infty, \qquad \ell = 0, \ldots, 3.$$

provided $f \in C^4([a,b])$.

We conclude with summing up the results in a theorem.

Theorem 6 *For piecewise polynomial interpolation on n equidistant points, we have the following error estimates.*

$$\begin{array}{rcllll}
||f - \tilde{p}||_\infty & \leq & C_k n^{-(k+1)}, & \tilde{p} \in C^0, & k\text{-th order Lagrange interpolation, } f \in C^{k+1}, \\
||f^{(\ell)} - \tilde{p}^{(\ell)}||_\infty & \leq & C_{\ell k} n^{-(2k+2-\ell)}, & \tilde{p} \in C^k, & k \text{ derivatives Hermite interpolation, } f \in C^{2k+2}, \\
||f^{(\ell)} - s^{(\ell)}||_\infty & \leq & C_\ell n^{-(4-\ell)}, & s \in C^2, & \text{cubic splines, } f \in C^4.
\end{array}$$

Final remark:

- The error estimates all require some smoohtness of f. For less regular functions the order of accuracy decreases. Suppose e.g. that $f \in C^s$ with $s < k+1$ for Lagrange interpolation. Then (4) must be replaced by

$$f(x) - p(x) = \frac{f^{(s)}(\xi) - p^{(s)}(\xi)}{n!}(x-y_1)(x-y_2)\cdots(x-y_s),$$

  where $\{y_j\}_{j=1}^s$ is any subset of $\{x_j\}_{j=1}^{k+1}$. By the same arguments that led up to (6), one obtains an order of accuracy of s, i.e. $||f - \tilde{p}_{k,m}||_\infty = \mathcal{O}(n^{-s})$. It therefore makes little sense to use high order interpolation ($k+1 > s$) for non-smooth f. This is a typical result also for other interpolation methods.

3.2 Better placed nodes

Another method of improving the bad convergence properties of high order equidistant interpolation is to use nodes that are not uniformly distributed in the interval. Given the error formula (4) a natural way to select the nodes is to find the points $\{x_j\} \subset [a,b]$ which minimize

$$\sup_{a \le x \le b} |x - x_1| \cdots |x - x_n|.$$

When the interval $[a,b] = [-1,1]$ this strategy gives the *Chebyshev* nodes,

$$x_j^c = -\cos(\theta_0 + (j-1)\Delta\theta), \qquad \theta_0 = \frac{\pi}{2n}, \qquad \Delta\theta = \frac{\pi}{n}, \qquad j = 1, \ldots, n.$$

The polynomials

$$T_n(x) = 2^{n-1}(x - x_1^c)\cdots(x - x_n^c)$$

are called the Chebyshev polynomials and they have the so-called min-max property

$$\max_{-1\le x\le 1} |2^{1-n} T_n(x)| = \min_{\omega \in \mathbb{P}_n^1} \max_{-1 \le x \le 1} |\omega(x)| = 2^{1-n},$$

where $\mathbb{P}_n^1$ are the polynomials of degree up to n with leading coefficient one. With these nodes one can for instance show the following error estimate. If $f \in C^s([-1,1])$, then

$$||f - p_n||_\infty \le C\, n^{-s} \log(1+n) ||f^{(s)}||_\infty.$$

Hence, if f is just continuously differentiable, $f \in C^1$, then Chebyshev interpolation always converges with n. The order of accuracy is only constrained by the regularity s of f.