Polynomial Interpolation: error analysis

KEY WORDS. interpolation, polynomial interpolation, error.

GOAL.

- To be able to plot the error function for visualization.
- To understand the error of polynomial interpolation.

1 The Error in Polynomial Interpolation

Let $p_n(x)$ be the polynomial of degree n interpolating the data

x	x_0	x_1	$\cdots$	x_n
y	y_0	y_1	$\cdots$	y_n

Assume that the data is given by a function $y = f(x)$ with the property that $y_i = f(x_i)$ and $x_i \in [a, b]$ for $i = 0, 1, \cdots, n$. The question that we consider here is: *how accurately does the polynomial $p_n(x)$ approximate the function $f(x)$ at any point x?*

2 An Example of the Error in Polynomial Interpolation

Let us look at an example of polynomial interpolation to gain some intuitive understanding of its error. Consider the function

$$f_1(x) = \frac{\sin(3x)}{1 + 3x}$$

defined on the interval $[a, b] = [0, 6]$. This function is plotted in Fig. 1.

The polynomial $p_5(x)$ of degree 5 interpolating this function at the six equally-spaced nodes:

$$x(i) = \frac{6i}{5}, \qquad i = 0, 1, \cdots, 5,$$

is plotted in Fig. 2, and the interpolation error $e_5(x) = f_1(x) - p_5(x)$ is plotted in Fig. 3. It is clear that the error can be quite large and the corresponding polynomial interpolation is not acceptable. However, if we use 13 equally-spaced nodes to interpolate $f_1(x)$ with a polynomial $p_{12}(x)$ of degree 12, then the error is plotted in Fig. 4. This new polynomial is much closer to the function $f_1(x)$ than $p_5(x)$. In other words, it might appear that functions can be better interpolated

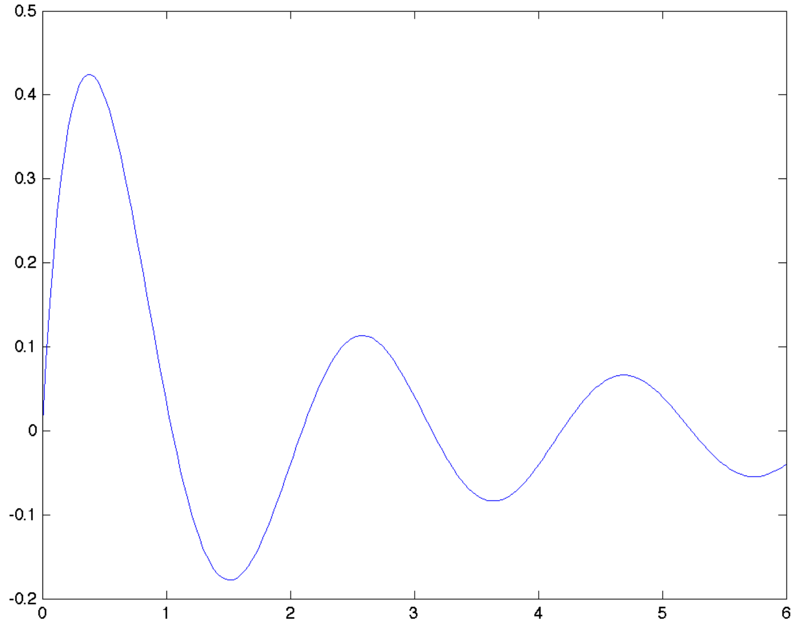

Figure 1: Plot of $f_1(x) = \sin(3x)/(1 + 3x)$.

by polynomials when more interpolation points are used.... but *this is not always the case* as we shall see later in this chapter.

The following MATLAB scripts were used to generate the figures. First, we need a MATLAB function to compute the coefficients in the Newton divided difference interpolating polynomial. This is given as follows:

```
function c=divdif(x_nodes,f_values)
%
% To compute the coefficients c(1),...,c(n) in the Newton form:
%  p(x) = c(1) + x(2) (x-x(1)) + ... + c(n) (x-x(1))...(x-x(n-1))
%
divdif_f= f_values;
n=length(x_nodes);
for i=2:n
  for j =n:-1:i
    divdif_f(j)=(divdif_f(j)-divdif_f(j-1)) ./(x_nodes(j)-x_nodes(j-i+1));
  end
  c(i)=divdif_f(i);
```

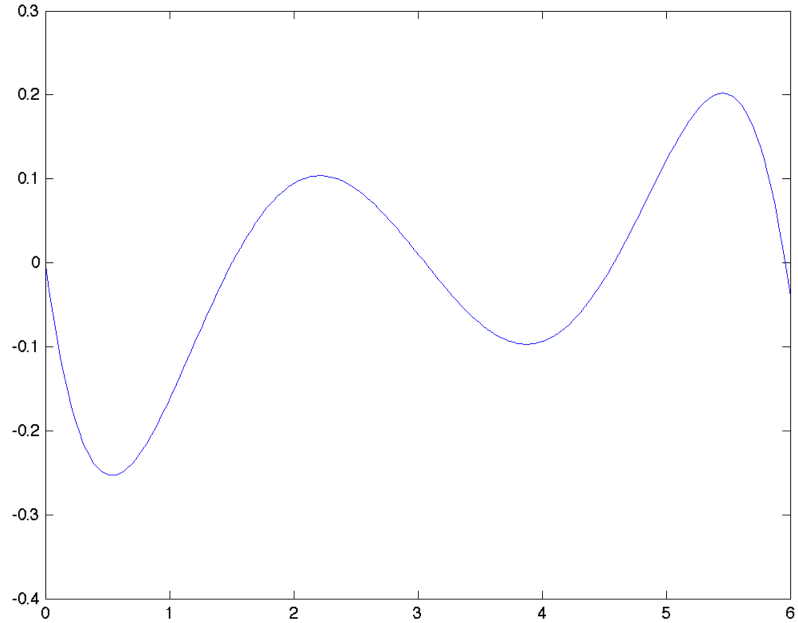

Figure 2: Plot of $p_5(x)$ from six equally-spaced nodes.

```
end
```

Next, we need to use nested multiplication to evaluate the Newton divided difference interpolating polynomial. The following MATLAB function will do this:

```
function pval=NestedM(c,x,z)
n=length(c);
pval = c(n)*ones(size(z));
for k=n-1:-1:1
  pval= (z-x(k)).*pval+c(k);
end
```

To plot the polynomial and the error, we use the following in MATLAB:

```
>> xint=linspace(0,6,13);
>> yint=sin(3*xint)./(1+3*xint);
>> cint=divdif(xint,yint);
```

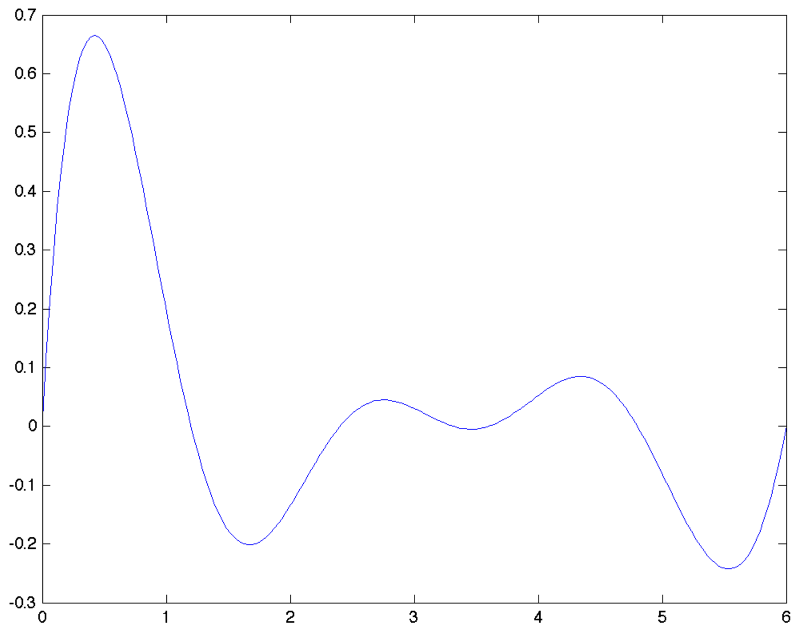

Figure 3: Plot of the error $e_5(x) = f_1(x) - p_5(x)$.

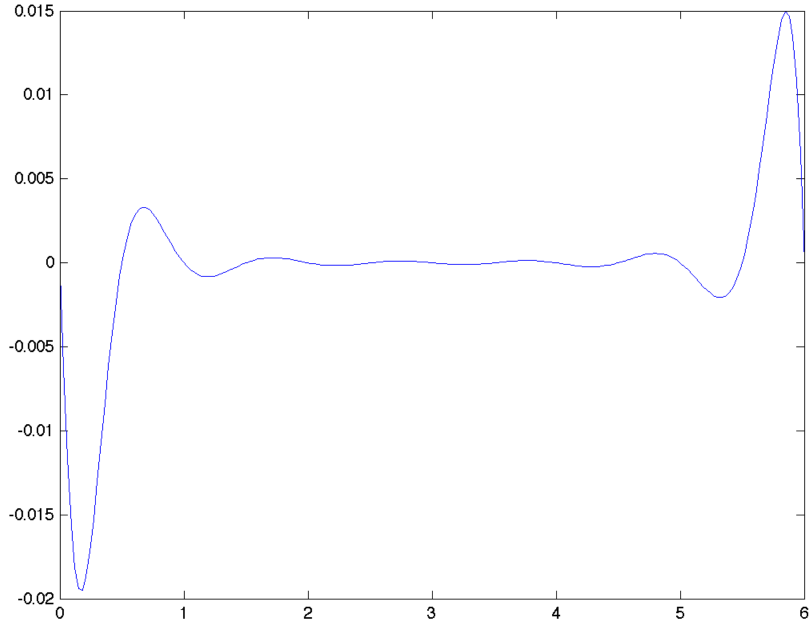

Figure 4: Plot of the error $e_{12}(x) = f_1(x) - p_{12}(x)$.

```
>> xn=linspace(0, 6, 200);
>> pval=NestedM(cint, xint, xn);
>> plot(xn, pval);
>> wn=sin(3*xn)./(1+3*xn);
>> plot(xn, wn-pval);
>> plot(xn, wn-pval, xn, pval);
```

- The first two lines define the interpolation data (x_i, f_i) for $i = 1, 2, \ldots, 13$, with $x_i = 0.5 * (i - 1)$.
- The third line computes the coefficients c_i of the Newton divided difference interpolating polynomial.
- The fourth line takes a sample of 200 points uniformly distributed on the interval $(0, 6)$; this is for plotting.
- The fifth line evaluates the Newton divided difference interpolating polynomial at the 200 sample points, and the values are saved in the vector pval().

- The sixth line plots the polynomial.
- The seventh line computes the value of the given function $f(x)$ at the 200 sampling points.
- The eighth line plots the interpolation error.
- The last line plots both the error function and the interpolating polynomial.

3 Error Estimates

Let the evaluation point be x and let all the nodes $\{x_i\}_{i=0}^n$ lie in a closed interval $[a, b]$. Then, as we shall prove shortly, if the function f has $n+1$ continuous derivatives on the interval $[a, b]$, the error expression takes the form

$$f(x) - p_n(x) = \frac{\omega_{n+1}(x)}{(n+1)!} f^{(n+1)}(\xi_x), \tag{1}$$

where

$$\omega_{n+1}(x) \equiv (x - x_0)(x - x_1) \cdots (x - x_n) = \prod_{j=0}^{n} (x - x_j),$$

and ξ_x is some (unknown) point in the interval $[a, b]$. The precise location of this point depends on $\{x_i\}_{i=0}^n$. Here $f^{(n+1)}(\xi_x)$ is the $(n+1)$st derivative of $f(x)$ evaluated at the point $x = \xi_x$.

To prove the desired error expression (1), note first that the result is trivially true when x is any node x_i since then both sides of the expression are zero. Assume that x does not equal to any node and consider the function $F(t)$ where

$$F(t) = f(t) - p_n(t) - c\omega_{n+1}(t),$$

and

$$c = \frac{[f(x) - p_n(x)]}{\omega_{n+1}(x)}.$$

Observe that c is well defined because $\omega_{n+1}(x) \neq 0$ since x is not a node. Note also that $F(x_i) = 0, \quad i = 0, \ldots, n$, and $F(x) = 0$. Thus $F(t)$ has at least $n+2$ distinct zeros in $[a, b]$. Now invoke Mean Value theorem which states that between any two zeros of F there must occur a zero of F'. Thus, F' has at least $n+1$ distinct zeros. By similar reasoning, F'' has at least n distinct zeros, and so on. Finally, it can be inferred that $F^{(n+1)}$ must have at least one zero. Let ξ_x be a zero of $F^{(n+1)}(t)$. Thus we have

$$0 = F^{(n+1)}(\xi_x) = f^{(n+1)}(\xi_x) - c(n+1)! = f^{(n+1)}(\xi_x) - \frac{(n+1)!}{\omega_{n+1}(x)}[f(x) - p_n(x)],$$

since $w_{n+1}^{(n+1)}(t) = (n+1)!$. The desired result (1) follows.

The following are some of the intrinsic properties of the interpolation error:

- For any value of i, the error is zero when $x = x_i$ because

$$\omega_{n+1}(x_i) = 0$$

(the interpolating conditions).

- The error is zero when the data f_i are measurements of a polynomial $f(x)$ of exact degree n because then the $(n+1)$st derivative,

$$f^{(n+1)}(\xi_x),$$

is identically zero. This is simply a statement of the uniqueness theorem of polynomial interpolation.

Taking absolute values in the interpolation error expression and maximizing both sides of the resulting inequality over $x \in [a,b]$, we obtain the polynomial interpolation error bound

$$\max_{x\in[a,b]} |f(x) - p_n(x)| \leq \max_{x\in[a,b]} |\omega_{n+1}(x)| \cdot \frac{\max_{x\in[a,b]} |f^{(n+1)}(x)|}{(n+1)!}. \tag{2}$$

Therefore, we are left with estimating the terms on the right-hand side of this inequality in order to characterize the error in polynomial interpolation.

4 Error Estimates for Uniformly Spaced Nodes

Consider the special case in which the points $\{x_i\}_{i=0}^n$ are equally-spaced on an interval $[a,b]$. Specifically, these points x_i are defined by

$$x_i = a + \frac{i}{n}(b-a), \quad i = 0, 1, \cdots, n.$$

We would like to derive an estimate for the following term appearing in (2):

$$W = \max_{x\in[a,b]} |\omega_{n+1}(x)|.$$

To this end, we make the change of variable

$$x = a + \frac{(b-a)s}{n}$$

to rewrite $\omega_{n+1}(x)$ as

$$\omega_{n+1}(x) = \prod_{j=0}^{n}(x - x_j) = \left(\frac{b-a}{n}\right)^{n+1} \prod_{j=0}^{n}(s-j),$$

where $s \in (0,n)$. Therefore, we have the following estimate:

$$W = \left(\frac{b-a}{n}\right)^{n+1} \max_{s\in(0,n)} \prod_{j=0}^{n} |s-j|. \tag{3}$$

For any given $s \in (0, n)$, let i be an integer such that $i < s < i + 1$. It follows that

$$\prod_{j=0}^{n} |s - j| = |(s - i)(s - i - 1)| \prod_{j=0}^{i-1} |s - j| \prod_{j=i+2}^{n} |s - j|, \tag{4}$$

and, since $s < i + 1$,

$$|(s - i)(s - i - 1)| \leq \frac{1}{4}, \qquad \text{(please show this)} \tag{5}$$

$$\prod_{j=0}^{i-1} |s - j| \leq \prod_{j=0}^{i-1} (i + 1 - j) \leq (i + 1)! \tag{6}$$

and, since $s > i$

$$\prod_{j=i+2}^{n} |s - j| \leq \prod_{j=i+2}^{n} (j - i) \leq (n - i)! \tag{7}$$

On substituting these three estimates, (5)–(7) into (4), we obtain

$$\prod_{j=0}^{n} |s - j| \leq \frac{1}{4} n! \tag{8}$$

On substituting (8) into (3), we obtain

$$W \leq \left(\frac{b - a}{n} \right)^{n+1} \frac{1}{4} n!,$$

and with this bound in (2), we obtain the following estimate:

$$\max_{x \in [a,b]} |f(x) - p_n(x)| \leq \left(\frac{b - a}{n} \right)^{n+1} \frac{1}{4(n + 1)} \max_{x \in [a,b]} |f^{(n+1)}(x)|. \tag{9}$$

Thus, if a function has ill-behaved higher derivatives, then the quality of the polynomial interpolation may actually decrease as the degree of the polynomial increases.

EXAMPLE 1: Consider $f(x) = \sin x$ and suppose values are known at three equally-spaced nodes x_0, x_1, and x_2. Then, from (1), the error in approximating the function $\sin x$ by $p_2(x)$ is

$$f(x) - p_2(x) = \frac{(x - x_0)(x - x_1)(x - x_2)}{3!} f^{(3)}(\xi_x) = \frac{1}{6} \omega_3(x) f^{(3)}(\xi_x).$$

If, for simplicity, we set $x_0 = -h$, $x_1 = 0$, and $x_2 = h$, then

$$|\omega_3(x)| = |x^3 - h^2 x|.$$

The maximum of this function between x_0 and x_2 occurs at $x = \pm h/\sqrt{3}$ and this maximum value is $2h^3/3\sqrt{3}$. (You may wish to verify this.) Moreover, since $f^{(3)}(x) = -\cos x$ (recall that $f(x) = \sin x$),

$$|f^{(3)}(\xi_x)| \leq 1,$$

we obtain

$$\max_{-h \leq x \leq h} |f(x) - p_2(x)| \leq \frac{\sqrt{3}}{27} h^3.$$

If, for example, we wish to obtain seven place accuracy using quadratic interpolation, we would have to choose h such that

$$\frac{\sqrt{3}}{27} h^3 < 5 \cdot 10^{-8}$$

Hence $h \approx 0.01$.

EXAMPLE 2: Determine the spacing h in a table of equally spaced values of the function

$$f(x) = \sqrt{x}$$

between 1 and 2, so that interpolation with a quadratic polynomial will yield an accuracy of 5×10^{-8}.

Solution: By assumption, the table will contain $f(x_i)$, with $x_i = 1 + ih$, $i = 0, 1, \ldots, N$, where $N = (2-1)/h = 1/h$. If $\bar{x} \in [x_{i-1}, x_{i+1}]$, then we approximate $f(\bar{x})$ by $p_2(\bar{x})$, where $p_2(x)$ is the quadratic polynomial which interpolates $f(x)$ at x_{i-1}, x_i, x_{i+1}. Then, from (1),

$$f(\bar{x}) - p_2(\bar{x}) = \frac{(\bar{x} - x_{i-1})(\bar{x} - x_i)(\bar{x} - x_{i+1})}{3!} f^{(3)}(\xi),$$

for some $\xi \in (x_{i-1}, x_{i+1})$. Since we do not know ξ, we can only estimate $f^{(3)}(\xi)$,

$$|f^{(3)}(\xi)| \leq \max_{1 \leq x \leq 2} |f^{(3)}(x)|.$$

Since

$$f^{(3)}(x) = \frac{3}{8} x^{-5/2},$$

it follows that $|f^{(3)}(\xi)| \leq \frac{3}{8}$. Further, from Example 1,

$$\max_{x \in [x_{i-1}, x_{i+1}]} |(\bar{x} - x_{i-1})(\bar{x} - x_i)(\bar{x} - x_{i+1})| \leq \frac{2h^3}{3\sqrt{3}}.$$

We are now assured that, for any $\bar{x} \in [1, 2]$,

$$|f(\bar{x}) - p_2(\bar{x})| \leq \frac{2h^3}{3\sqrt{3}} \times \frac{3}{8} \times \frac{1}{3!} = \frac{h^3}{24\sqrt{3}},$$

if $p_2(x)$ is chosen as the quadratic polynomial which interpolates $f(x) = \sqrt{x}$ at the three tabular points nearest $\bar{x}$.

For an accuracy of 5×10^{-8}, we must choose h so that

$$\frac{h^3}{24\sqrt{3}} < 5 \times 10^{-8},$$

giving $h \approx 0.01028$ or $N \approx 79$.