Interpolation and Approximation: Lagrange Interpolation

Martin Licht

UC San Diego

Winter Quarter 2021

Lagrange Interpolation

Given a function $f : [a, b] \to \mathbb{R}$ over some interval $[a, b]$, we would like to approximate f by a polynomial.

How do we find a good polynomial?

We have already one example, namely the Taylor polynomial around a point a:

$$T_a^m f(x) = \sum_{k=0}^{m} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

Note that this can be written as

$$T_a^m f(x) = \sum_{k=0}^{m} c_k (x-a)^k,$$

where

$$c_k = \frac{f^{(k)}(a)}{k!}.$$

Evidently, we construct the Taylor polynomial by evaluating f and its derivatives at a particular point $a \in \mathbb{R}$.

We recall some representations of the error:

Theorem

Let $f : \mathbb{R} \to \mathbb{R}$ have continuous derivatives up to order $m+1$. Then

- *We have*

$$R_a^m f(x) = \int_a^x \frac{f^{(m+1)}(t)}{m!}(t-a)^m \, \mathrm{d}t.$$

- *For every $x \in \mathbb{R}$ there exists ξ_x in the closed interval between a and x with*

$$R_a^m f(x) = \frac{f^{(m+1)}(\xi_x)}{(m+1)!}(x-a)^{m+1}.$$

- *For every $x \in \mathbb{R}$ there exists ξ_x in the closed interval between a and x with*

$$R_a^m f(x) = \frac{f^{(m+1)}(\xi_x)}{m!}(x-\xi)^m (x-a).$$

From each of those representations of the error we can derive

$$\left|f(x) - T_a^m f(x)\right| = \left|R_a^m f(x)\right| \leq \frac{1}{m!} \max_{\xi \in I} \left|f^{(m+1)}(\xi)\right| \cdot |x - a|^{m+1}.$$

or even

$$\left|f(x) - T_a^m f(x)\right| = \left|R_a^m f(x)\right| \leq \frac{1}{(m+1)!} \max_{\xi \in I} \left|f^{(m+1)}(\xi)\right| \cdot |x - a|^{m+1}.$$

where I is the interval between a and x.

Lagrange Interpolation

The basic principle of polynomial interpolation is that we "take measurements" of f by looking at the values of the function (and its derivatives) at certain points. We then construct a polynomial that satisfies the same measurements.

In the case of the Taylor polynomial, we have a single number $x_0 \in \mathbb{R}$ and take the derivatives up to order m, to construct a degree m polynomial $p(x)$ with

$$p(x_0) = f(x_0), \quad p'(x_0) = f'(x_0), \quad p''(x_0) = f''(x_0), \quad \dots \quad p^{(m)}(x_0) = f^{(m)}(x_0).$$

A different way of interpolating a function is known as Lagrange interpolation.

In the case of Lagrange interpolation, we have m different numbers $x_0, x_1, \dots, x_m \in \mathbb{R}$ and take function evaluations up to order m, to construct a degree m polynomial $p(x)$ with

$$p(x_0) = f(x_0), \quad p(x_1) = f(x_1), \quad p(x_2) = f(x_2), \quad \dots \quad p(x_m) = f(x_m).$$

Example

Suppose we have got points $x_0, x_1, \ldots, x_m$ and values

$$y_0 = f(x_0), \quad y_1 = f(x_1), \quad \ldots \quad y_m = f(x_m)$$

of some function f that is otherwise unknown. We want to reconstruct a polynomial that attains the same function values as f. For the sake of overview, we put this into a table:

x	x_0	x_1	$\ldots$	x_m
y	y_0	y_1	$\ldots$	y_m

For this example, let us consider the case $m = 2$ and

$$\begin{aligned} x_0 &= -1, \quad x_1 = 0, \quad x_2 = 1, \\ y_0 &= 6, \quad y_1 = 2, \quad y_2 = 4. \end{aligned}$$

Example

The table is

x	-1	0	1
y	6	2	4

We search for a polynomial p of degree $m = 2$ such that

$$p(-1) = 6, \quad p(0) = 2, \quad p(1) = 4.$$

The solution is the polynomial

$$p(x) = 2 - x + 3x^2.$$

In these notes, we describe different ways to computing and representing such polynomials.

Monomial Basis

Suppose we have pairwise different points $x_0, x_1, \dots, x_m$ and that we search for the coefficients $a_0, a_1, \dots, a_m$ of a polynomial

$$p(x) = a_0 + a_1 x + \cdots + a_m x_m$$

such that for some given values $y_0, y_1, \dots, y_m$ we have

$$p(x_0) = y_0, \quad p(x_1) = y_1, \quad \dots \quad p(x_m) = y_m.$$

That is, we search for m unknown variables $a_0, a_1, \dots, a_m \in \mathbb{R}$ such that the m constraints given by the point evaluations are satisfied. This translates into a linear system of equations

$$\begin{aligned}
a_0 + a_1 x_0 + a_2 x_0^2 + \cdots + a_m x_0^m &= y_0, \\
a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_m x_1^m &= y_1, \\
a_0 + a_1 x_2 + a_2 x_2^2 + \cdots + a_m x_2^m &= y_2, \\
&\vdots \\
a_0 + a_1 x_m + a_2 x_m^2 + \cdots + a_m x_m^m &= y_m.
\end{aligned}$$

We can rewrite this in matrix notation as

$$\begin{pmatrix} 1 & x_0 & x_0^2 & \dots & x_0^m \\ 1 & x_1 & x_1^2 & \dots & x_1^m \\ 1 & x_2 & x_2^2 & \dots & x_2^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \dots & x_m^m \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

The matrix in that system called the ***Vandermonde matrix*** associated to the points $x_0, x_1, \dots, x_m$. We would like to understand the linear system of equations has got a solution, and for that purpose the Vandermonde matrix.

Theorem

The determinant of the Vandermonde matrix V is

$$\det(V) = \prod_{0 \le i < j \le m} (x_j - x_i).$$

Proof.

For the proof we use elementary properties of determinants. Let $x_0, x_1, \dots, x_m \in \mathbb{R}$ be pairwise different. Since the determinant is invariant under row additions and subtractions, we get the identity

$$\det \begin{pmatrix} 1 & x_0 & x_0^2 & \dots & x_0^m \\ 1 & x_1 & x_1^2 & \dots & x_1^m \\ 1 & x_2 & x_2^2 & \dots & x_2^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \dots & x_m^m \end{pmatrix} = \det \begin{pmatrix} 1 & x_0 & x_0^2 & \dots & x_0^m \\ 0 & x_1 - x_0 & x_1^2 - x_0^2 & \dots & x_1^m - x_0^m \\ 0 & x_2 - x_0 & x_2^2 - x_0^2 & \dots & x_2^m - x_0^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & x_m - x_0 & x_m^2 - x_0^2 & \dots & x_m^m - x_0^m \end{pmatrix}$$

Similarly, the determinat is invariant under additions of columns. We perform a number of column substractions: we subtract x_0-times the m-th column from the $(m+1)$-th column, subtract x_0-times the $(m-1)$-th column from the m-th column, subtract x_0-times the $(m-2)$-th column from the $(m-1)$-th column, and and so on, until we have subtracted x_0-times the first column from the second column. □

Proof.

Consequently, we end up with the determinant

$$\det \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & x_1 - x_0 & (x_1 - x_0)x_1 & \dots & (x_1 - x_0)x_1^{m-1} \\ 0 & x_2 - x_0 & (x_2 - x_0)x_2 & \dots & (x_2 - x_0)x_2^{m-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & x_m - x_0 & (x_m - x_0)x_m & \dots & (x_m - x_0)x_m^{m-1} \end{pmatrix}$$

The rows of this determinant have the common factors

$$(x_1 - x_0), \quad (x_2 - x_0), \quad \dots \quad (x_m - x_0).$$

□

Proof.

We can extract these common factors from the determinant and get the value

$$\prod_{i=1}^{m}(x_i - x_0) \cdot \det \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & x_1 & \dots & x_1^{m-1} \\ 0 & 1 & x_2 & \dots & x_2^{m-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & x_m & \dots & x_m^{m-1} \end{pmatrix}$$

$$= \prod_{i=1}^{m}(x_i - x_0) \cdot \det \begin{pmatrix} 1 & x_1 & \dots & x_1^{m-1} \\ 1 & x_2 & \dots & x_2^{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & \dots & x_m^{m-1} \end{pmatrix}$$

The last term is the determinat of the Vandermonde matrix for the points $x_1, \dots, x_m$. □

Proof.

We can repeat this calculation recursively until we only need to compute the determinant of the Vandermonde matrix for the single point x_0, which is just equals 1. Working up from there, the determinant becomes

$$\prod_{i=1}^{m}(x_i - x_0) \cdot \prod_{1\leq i<j\leq m}(x_j - x_i) = \prod_{0\leq i<j\leq m}(x_j - x_i)$$

This completes the proof. □

In particular, since $x_0, x_1, \ldots, x_m$ are pairwise different, the determinant of the Vandermonde matrix is non-zero, and hence that the matrix is invertible. We conclude that the interpolation problem has a got a unique solution.

Theorem

Given pairwise distinct points $x_0, x_1, \ldots, x_m \in \mathbb{R}$ and values $y_0, y_1, \ldots, y_m \in \mathbb{R}$, there exists a unique polynomial p of degree m such that

$$p(x_0) = f(x_0), \quad p(x_1) = f(x_1), \quad p(x_2) = f(x_2), \quad \ldots \quad p(x_m) = f(x_m).$$

The polynomials of degree m are a vector space of dimension $m+1$, with a basis being the monomials up to order m:

$$1, \quad x, \quad x^2, \quad \ldots \quad x^m,$$

In particular, if we express the interpolation problem using the monomial basis, then the basis does not depend on the interpolation points $x_0, x_1, \ldots, x_m$.

However, the Vandermonde matrix in the formulation has several disadvantageous properties, e.g., it is very dense.

Example

Consider again the quadratic interpolation problem with the following table:

x	-1	0	1
y	6	2	4

The solution is

$$\begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \\ 4 \end{pmatrix}.$$

We check that the determinant of Vandermonde matrix is

$$\det \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} = (1)(2)(1) = 2.$$

Example

The inverse of that Vandermonde matrix is

$$\begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}^{-1} = \frac{1}{2}\begin{pmatrix} 0 & 2 & 0 \\ -1 & 0 & 1 \\ 1 & -2 & 1 \end{pmatrix},$$

and we readily check that

$$\frac{1}{2}\begin{pmatrix} 0 & 2 & 0 \\ -1 & 0 & 1 \\ 1 & -2 & 1 \end{pmatrix}\begin{pmatrix} 6 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix},$$

which is precisely the coefficients of the solution $p(x) = 2 - x + 3x^2$.

Newton Polynomials

We pose the same interpolation but with a different basis. This time, the basis incorporates the interpolation points $x_0, x_1, \ldots, x_m \in \mathbb{R}$. We define the Newton polynomials

$$\begin{aligned}
p_0(x) &= 1 \\
p_1(x) &= (x - x_0) \\
p_2(x) &= (x - x_0)(x - x_1) \\
p_3(x) &= (x - x_0)(x - x_1)(x - x_2) \\
&\vdots \\
p_m(x) &= (x - x_0)(x - x_1)(x - x_2) \cdot \cdots \cdot (x - x_{m-1})
\end{aligned}$$

So we have the form

$$p_k(x) = \prod_{i=0}^{k-1} (x - x_k) = (x - x_0)(x - x_1) \cdot \cdots \cdot (x - x_{k-1}).$$

Consequently,

$$p_k(x_0) = \cdots = p_k(x_{k-1}) = 0.$$

Newton Polynomials

Using this basis lets us formulate the interpolation problem in a simplified manner. Using the Newton polynomials, we search coefficients $a_0, a_1, \ldots, a_m \in \mathbb{R}$ such that

$$\begin{aligned}
a_0p_0(x_0) + a_1p_1(x_0) + a_2p_2(x_0) + \cdots + a_mp_m(x_0) &= y_0, \\
a_0p_0(x_1) + a_1p_1(x_1) + a_2p_2(x_1) + \cdots + a_mp_m(x_1) &= y_1, \\
a_0p_0(x_2) + a_1p_1(x_2) + a_2p_2(x_2) + \cdots + a_mp_m(x_2) &= y_2, \\
&\vdots \\
a_0p_0(x_m) + a_1p_1(x_m) + a_2p_2(x_m) + \cdots + a_mp_m(x_m) &= y_m.
\end{aligned}$$

This can be written in matrix notation as

$$\begin{pmatrix} p_0(x_0) & 0 & 0 & \ldots & 0 \\ p_0(x_1) & p_1(x_1) & 0 & \ldots & 0 \\ p_0(x_2) & p_1(x_2) & p_2(x_2) & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_0(x_m) & p_1(x_m) & p_2(x_m) & \ldots & p_m(x_m) \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

Newton Polynomials

The coefficients in that matrix have an explicit form:

$$p_i(x_j) = (x_j - x_0)(x_j - x_1) \cdot \dots \cdot (x_j - x_{i-1}).$$

The system of equations is invertible because the matrix triangular and the diagonal entries are non-zero, hence the matrix is invertible. We recall that such systems can be solved via forward substitution.

$$\begin{aligned}
a_0 &= y_0, \\
a_1 &= p_1(x_1)^{-1}(y_1 - a_0 p_0(x_1)), \\
a_2 &= p_2(x_2)^{-1}(y_2 - a_1 p_1(x_2) - a_0 p_0(x_1)), \\
&\vdots \\
a_m &= p_m(x_m)^{-1}(y_m - a_{m-1}p_{m-1}(x_m) - \dots - a_1 p_1(x_m) - a_0 p_0(x_m)).
\end{aligned}$$

In particular, the coefficient a_i only depends on the points $x_0, x_1, \dots, x_i$. This shows the following fact: we can keep adding points and incrementally construct the coefficients.

Example

Once more, consider the quadratic interpolation problem with the following table:

x	-1	0	1
y	6	2	4

The linear system corresponding to the Newton polynomials is

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \\ 4 \end{pmatrix}.$$

The inverse of the matrix is

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 2 \end{pmatrix}^{-1} = \frac{1}{2} \begin{pmatrix} 2 & 0 & 0 \\ 2 & 2 & 0 \\ 1 & -2 & 1 \end{pmatrix}.$$

Example

We readily check that

$$\frac{1}{2}\begin{pmatrix} 2 & 0 & 0 \\ 2 & 2 & 0 \\ 1 & -2 & 1 \end{pmatrix}\begin{pmatrix} 6 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 6 \\ -4 \\ 3 \end{pmatrix}.$$

This means that the polynomial can be represented as

$$\begin{aligned} p(x) &= 6 + (-4)(x - x_0) + 3(x - x_1)(x - x_0) \\ &= 6 + (-4)(x + 1) + 3(x)(x + 1) \\ &= 6 + ((-4)x + (-4)) + (3x^2 + 3x) \\ &= 6 - 4x - 4 + 3x^2 + 3x \\ &= 3x^2 - x + 2. \end{aligned}$$

Example

Instead of computing the inverse matrix, forward substitution is a quick way of getting the coefficients. We get

$$\begin{aligned} a_0 &= 6, \\ a_1 &= 2 - 1a_0 = 2 - 6 = -4, \\ 2a_2 &= 4 - 1a_0 - 2a_1 = 4 - 6 - 2(-4) = 6 \implies a_2 = 3. \end{aligned}$$

Suppose we augment our table of input/output values with

$$x_3 = 2, \quad y_3 = 6.$$

We have $p_3(x) = (x+1)(x)(x-1)$ being a cubic polynomial that vanishes on the first points x_0, x_1, x_2. The linear system of equations is extended by an additional line

$$a_0 p_0(x_3) + a_1 p_1(x_3) + a_2 p_2(x_3) + a_3 p_3(x_3) = y_3.$$

Example

We reuse the coefficients a_0, a_1, a_2 found previously and get

$$a_0 + 3a_1 + 6a_2 + 6a_3 = y_3,$$

so we calculate

$$\begin{aligned} 6a_3 &= y_3 - a_0 - 3a_1 - 6a_2 \\ &= y_3 - 6 - 3(-4) - 6(6) = y_3 - 6 + 12 - 6(6) \\ \implies \qquad a_3 &= \frac{1}{6}y_3 - 1 + 2 - 6 = 4. \end{aligned}$$

Consequently, the cubic polynomial $q(x)$ solving the interpolation problem is

$$\begin{aligned} q(x) &= 6 + (-4)(x - x_0) + 3(x - x_1)(x - x_0) + 4(x - x_2)(x - x_1)(x - x_0) \\ &= 6 + (-4)(x + 1) + 3(x)(x + 1) + 4(x - 1)(x)(x + 1). \end{aligned}$$

Lagrange Interpolation

Monomial Basis

Newton Polynomials

Lagrange Polynomials

Error Analysis for Lagrange Polynomials

Lagrange Polynomials

We define the Lagrange polynomials as follows:

$$l_k(x) = \prod_{\substack{0 \le i \le m \\ i \ne k}} \frac{x - x_i}{x_k - x_i}.$$

The Lagrange polynomials are polynomials of degree m: they are the product of m different factors each of which has the form $(x - x_i)(x_k - x_i)$.
Furthermore, the Lagrange polynomials satisfy the property

$$l_k(x_j) = \prod_{\substack{0 \le i \le m \\ i \ne k}} \frac{x_j - x_i}{x_k - x_i} = \begin{cases} 1 & \text{if } k = j \\ 0 & \text{if } k \ne j \end{cases}$$

The linear system associated with the interpolation problem ...

$$\begin{aligned}
a_0 l_0(x_0) + a_1 l_1(x_0) + a_2 l_2(x_0) + \cdots + a_m l_m(x_0) &= y_0, \\
a_0 l_0(x_1) + a_1 l_1(x_1) + a_2 l_2(x_1) + \cdots + a_m l_m(x_1) &= y_1, \\
a_0 l_0(x_2) + a_1 l_1(x_2) + a_2 l_2(x_2) + \cdots + a_m l_m(x_2) &= y_2, \\
&\vdots \\
a_0 l_0(x_m) + a_1 l_1(x_m) + a_2 l_2(x_m) + \cdots + a_m l_m(x_m) &= y_m.
\end{aligned}$$

Lagrange Polynomials

... can thus be rewritten in matrix notation as

$$\begin{pmatrix} l_0(x_0) & 0 & 0 & \dots & 0 \\ 0 & l_1(x_1) & 0 & \dots & 0 \\ 0 & 0 & l_2(x_2) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & l_m(x_m) \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

But this is just the identity matrix! In other words, $a_i = y_i$ is the solution. Hence

$$p(x) = y_0 l_0(x) + y_1 l_1(x) + \cdots + y_m l_m(x).$$

is the unique polynomial satisfying the Lagrange interpolation property.

Example

One more time, we consider the quadratic interpolation problem

x	-1	0	1
y	6	2	4

We only need to calculate the Lagrange polynomials:

$$l_0(x) = \frac{x - x_1}{x_0 - x_1}\frac{x - x_2}{x_0 - x_2} = \frac{x}{-1}\frac{x-1}{-2},$$
$$l_1(x) = \frac{x - x_0}{x_1 - x_0}\frac{x - x_2}{x_1 - x_2} = \frac{x+1}{1}\frac{x-1}{-1},$$
$$l_2(x) = \frac{x - x_0}{x_2 - x_0}\frac{x - x_1}{x_2 - x_1} = \frac{x+1}{2}\frac{x}{1}.$$

The solution is the quadratic polynomial

$$p(x) = 6l_0(x) + 2l_1(x) + 4l_2(x).$$

Error Analysis for Lagrange Polynomials

We have seen different ways of computing and writing the solution to the interpolation problem. Suppose we have a found a polynomial $p(x)$ of degree m such that

$$p(x_0) = y_0, \quad p(x_1) = y_1, \quad \dots \quad p(x_m) = y_m.$$

In many cases, the output values y_i are values assumed by some function f,

$$f(x_0) = y_0, \quad f(x_1) = y_1, \quad \dots \quad f(x_m) = y_m.$$

Hence the polynomial p ***interpolates*** the function f at the points $x_0, x_1, \dots, x_m$ and can be seen as an approximation of f.

We ask whether the interpolation polynomial is a good approximation of f. This is answered by the following result.

Theorem

Let $[a,b]$ *containing the pairwise distinct points* $x_0, x_1, \ldots, x_m$. *Let* $f : [a,b] \to \mathbb{R}$ *be a function with* $m+1$ *continuous derivatives over* $[a,b]$, *and let* $p(x)$ *be the m-th degree polynomial which solves*

$$p(x_0) = y_0, \quad p(x_1) = y_1, \quad \ldots \quad p(x_m) = y_m.$$

Then for all $x \in [a,b]$ *there exists* $\xi(x) \in [a,b]$ *such that*

$$f(x) - p(x) = \frac{\omega(x)}{(m+1)!} f^{(m+1)}(\xi(x)), \quad \omega_{m+1}(x) = \prod_{i=0}^{m}(x - x_i).$$

Proof.

Let $x \in [a, b]$. The result is certainly true if x is one of the data points $x_0, x_1, \ldots, x_m$. Consider the case that x is any other point in $[a, b]$, so $\omega(x) \neq 0$ We define the function

$$g_x : [a, b] \to \mathbb{R}, \quad t \mapsto f(x) - p(x) - \frac{f(x) - p(x)}{\omega(x)} \omega(t)$$

We observe that g_x has at least $m + 2$ zeroes over $[a, b]$, namely the pairwise distinct points $x, x_0, x_1, \ldots, x_m$. By Rolle's theorem, g'_x has at least $m + 1$ zeroes over $[a, b]$, since between two consecutive zeroes of g_x there must be a zero of g'_x. □

Proof.

By similar reasoning, g_x'' has at least m zeroes, and repeating this argument shows that the k-th derivative of g_x has at least $m+2-k$ zeroes. In particular, $g_x^{(m+1)}$ has at least one zero, which we call $\xi_x \in [a,b]$.

Hence

$$\begin{aligned} 0 &= g_{\xi_x}^{(m+1)} \\ &= f^{(m+1)}(x) - p^{(m+1)}(x) - \frac{f(x)-p(x)}{\omega(x)}(m+1)! \\ &= f^{(m+1)}(x) - \frac{f(x)-p(x)}{\omega(x)}(m+1)!. \end{aligned}$$

Rearranging gives the desired result. □

Error Analysis for Lagrange Polynomials

As a consequence, if the function f has $m+1$ continuous derivatives over $[a,b]$, then for any $x \in [a,b]$ we can estimate

$$|f(x) - p(x)| \leq \frac{1}{(m+1)!} \cdot \prod_{i=0}^{m} |x - x_i| \cdot \max_{\xi \in [a,b]} \left| f^{(m+1)}(\xi) \right| .$$