Interpolation

M.Sc. in Mathematical Modelling & Scientific Computing,
Practical Numerical Analysis

Michaelmas Term 2020, Lecture 1

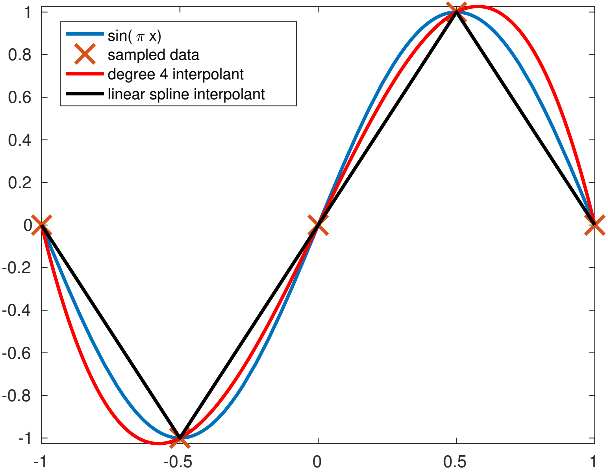
What is Interpolation?

Canonical interpolation problem: given a set of nodes $\mathbf{x}_i$, $0 \leq i \leq n$ and data at those nodes $f(\mathbf{x}_i)$, construct a function $p(\mathbf{x})$ such that

$$p(\mathbf{x}_i) = f(\mathbf{x}_i)$$

for $0 \leq i \leq n$.

Examples



sin(π x)
sampled data
degree 4 interpolant
linear spline interpolant
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1
-0.5
0
0.5
1

Types of Interpolant

There are lots of functions, $p(\mathbf{x})$ satisfying the interpolation conditions so we restrict ourselves to certain classes of functions, for example

- polynomials
- piecewise polynomials (splines in 1D)
- linear combinations of radial basis functions
- rational functions
- ...

Why Interpolate?

- The values $f(\mathbf{x}_i)$ may come from measuring data so we might want to find approximate values of $f(\mathbf{x})$ when $\mathbf{x}$ is not a measurement point.
- I might know the function $f(\mathbf{x})$ but it might be very complicated so I might want to represent it more simply to evaluate it more easily.
- In particular if $p(\mathbf{x}) \approx f(\mathbf{x})$ then we might expect roots of $p(\mathbf{x})$ to be good approximations of roots of $f(\mathbf{x})$.
- I might want to know $\int f(\mathbf{x})\mathrm{d}\mathbf{x}$ — a good approximation should be $\int p(\mathbf{x})\mathrm{d}\mathbf{x}$.

Lagrange Interpolation

To begin we consider interpolation in 1D by polynomials. The problem is now:

Given the distinct points x_0, x_1, ..., x_n and data values $f(x_0)$, $f(x_1)$, ..., $f(x_n)$, construct the polynomial $p_n(x)$, of degree at most n, such that $p_n(x_i) = f(x_i)$ for $i = 0, 1, \ldots, n$.

The polynomial $p_n(x)$ is known as the *Lagrange interpolation polynomial* (or simply the *Lagrange interpolant*).

Questions about the Lagrange Interpolant

There are many questions we can ask about the Lagrange interpolant, for example:

- Does p_n exist?
- Is p_n unique?
- Does p_n converge to f as $n \to \infty$?
- How should we store, compute and evaluate p_n?

Existence of the Lagrange Interpolant

Define

$$L_{n,k}(x) = \prod_{\substack{j=0 \\ j \neq k}}^{n} \frac{x - x_j}{x_k - x_j} \, .$$

Clearly $L_{n,k}(x)$ is a polynomial of degree n satisfying $L_{n,k}(x_i) = \delta_{i,k}$.

We can write

$$p_n(x) = \sum_{k=0}^{n} f(x_k) L_{n,k}(x)$$

and $p_n(x)$ is a polynomial of degree at most n. The interpolation conditions are satisfied because

$$p_n(x_i) = \sum_{k=0}^{n} f(x_k) L_{n,k}(x_i) = \sum_{k=0}^{n} f(x_k) \delta_{i,k} = f(x_i)$$

for $i = 0, 1, \ldots, n$. Hence $p_n(x)$ is a Lagrange interpolant and we have demonstrated existence of the Lagrange interpolant.

Uniqueness of the Lagrange Interpolant

Suppose that $p_n(x)$ and $q_n(x)$ are polynomials of degree at most n and that both satisfy the interpolation conditions. Then $(p_n - q_n)(x)$ is also a polynomial of degree at most n and we have

$$(p_n - q_n)(x_i) \quad = \quad 0$$

for $i = 0, 1, \ldots, n$ (from the interpolation conditions). Thus $(p_n - q_n)(x)$ is a polynomial of degree at most n with at least $n + 1$ distinct roots and so the fundamental theorem of algebra tells us that $(p_n - q_n)(x) \equiv 0$. Hence $p_n(x) \equiv q_n(x)$ and the Lagrange interpolant is unique.

This is useful because it tells us that if we find *a* polynomial of degree at most n satisfying the interpolation conditions, then we have found *the* Lagrange interpolant.

Error in the Lagrange Interpolant

Theorem 1 (Error Representation)

Suppose $f \in C^{n+1}[a,b]$ *(where* $x_i \in [a,b]$, $0 \le i \le n$*). Then for all* $x \in [a,b]$*, there exists* $\xi = \xi(x) \in (a,b)$ *such that*

$$f(x) - p_n(x) \;=\; \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{k=0}^{n} (x - x_k)\ .$$

In particular

$$|f(x) - p_n(x)| \;\le\; \frac{M_n}{(n+1)!} \prod_{k=0}^{n} (x - x_k)\ ,$$

where

$$M_n \;=\; \max_{\zeta \in [a,b]} |f^{(n+1)}(\zeta)|\ .$$

Error in the Lagrange Interpolant

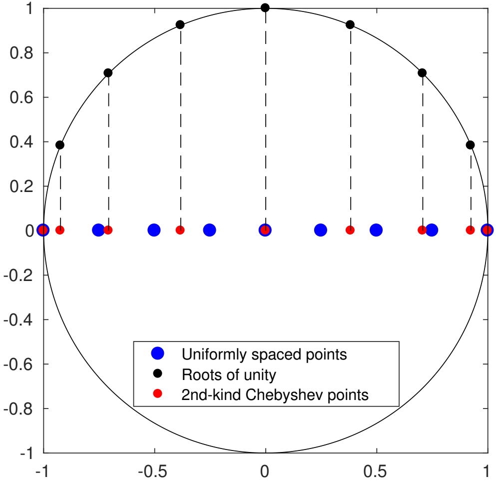
Theorem 1 tells us that the size of the error in the polynomial interpolant depends on both the size of the derivatives of the function we wish to interpolate and the grid points.

We consider two types of grid on $[-1, 1]$:

- Uniformly spaced grids $x_i = -1 + 2i/n$, $0 \leq i \leq n$
- Chebyshev grids $x_i = \cos(i\pi/n)$, $0 \leq i \leq n$ (these are known as second-kind Chebyshev points and are equally spaced on the unit circle)

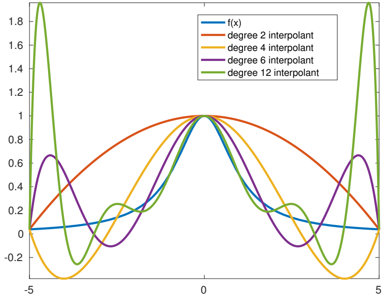
(note that these can be shifted and scaled to other intervals as appropriate).

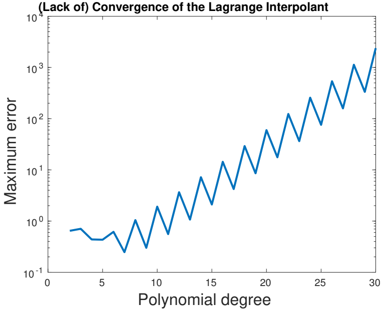
Grid Points


1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1
-0.5
0
0.5
1
Uniformly spaced points
Roots of unity
2nd-kind Chebyshev points

Example

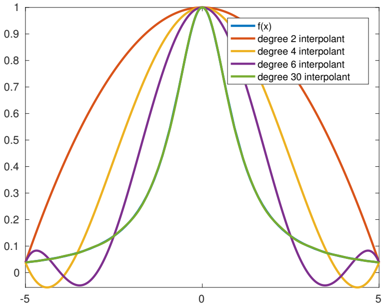
We interpolate the Runge function $f(x) = 1/(1 + x^2)$ on the interval $[-5, 5]$ using uniformly spaced points.

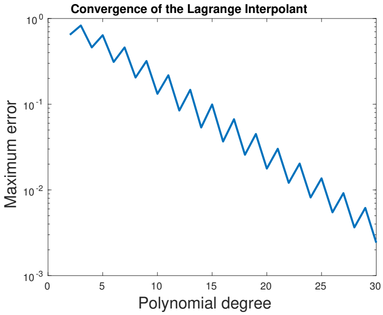




Example

We interpolate the Runge function $f(x) = 1/(1+x^2)$ on the interval $[-5, 5]$ using Chebyshev points.





Error in the Lagrange Interpolant

If we use a uniform grid on $[-1, 1]$, Theorem 1 tells us that, for $f \in C^{n+1}[-1, 1]$

$$|f(x) - p_n(x)| \quad \leq \quad \frac{M_n}{(n+1)!} \prod_{k=0}^{n} (x - x_k) \ ,$$

where

$$M_n \quad = \quad \max_{\zeta \in [-1,1]} |f^{(n+1)}(\zeta)| \ .$$

Hence

$$\max_{x \in [-1,1]} |f(x) - p_n(x)| \quad \leq \quad \frac{M_n}{(n+1)!} \max_{x \in [-1,1]} \left| \prod_{k=0}^{n} (x - x_k) \right| \ .$$

It can be shown that with a uniform grid of size $h = 2/n$

$$\max_{x \in [-1,1]} \left| \prod_{k=0}^{n} (x - x_k) \right| \quad \leq \quad n! h^{n+1} \ .$$

Error in the Lagrange Interpolant

Hence we have

$$\max_{x\in[-1,1]} |f(x) - p_n(x)| \;\; \leq \;\; \frac{M_n h^{n+1}}{n+1} \,.$$

This means that if M_n grows faster than $h^{n+1}/(n+1)$ decays, the polynomial interpolant on a uniform mesh can diverge. This is known as the Runge phenomenon.

Convergence Theorems

Theorem 2
On equispaced grids, polynomial interpolation can diverge even if f is analytic.

Theorem 3
On Chebyshev grids, if f is analytic, then $\|f - p_n\| = \mathcal{O}(\rho^{-n})$ where ρ depends on the position of the poles. (Geometric convergence)

Theorem 4
On Chebyshev grids, if $f^{(\nu)}$ with $\nu \geq 1$ has bounded variation V, then $\|f - p_n\| = \mathcal{O}(Vn^{-\nu})$.

Representation of the Lagrange Interpolant

Since the Lagrange interpolant is a polynomial of degree at most n it seems natural to write it as

$$p_n(x) \quad = \quad a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n \ , \qquad (1)$$

i.e. as a linear combination of monomial basis functions. Once we know the coefficients a_k, we can evaluate $p_n(x)$ using this expression and this requires $\mathcal{O}(n^2)$ operations. Alternatively we can use the *Horner scheme*.

The Horner Scheme

The Horner scheme makes use of the fact that we can rewrite Equation (1) as

$$p_n(x) \quad = \quad a_0 + x(a_1 + x(a_2 + \ldots + x(a_{n-1} + a_n x) \ldots)).$$

Then we set

$$\begin{aligned} b_n &= a_n \\ b_{n-1} &= a_{n-1} + b_n x \\ &\vdots \\ b_0 &= a_0 + b_1 x \\ p_n(x) &= b_0 \ . \end{aligned}$$

This is much more efficient than the naïve evaluation of $p_n(x)$ since it requires only $\mathcal{O}(n)$ flops.

Calculating the Coefficients — Vandermonde Matrices

If we write

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

then the interpolation conditions become

$$p_n(x_i) \;=\; a_0 + a_1 x_i + a_2 x_i^2 + \ldots + a_n x_i^n \;=\; f(x_i) \;,$$

for $i = 0, 1, \ldots, n$. We can represent this as a linear system of the form $V\mathbf{a} = \mathbf{f}$, i.e.

$$\begin{pmatrix} 1 & x_0 & \cdots & x_0^n \\ 1 & x_1 & \cdots & x_1^n \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \cdots & x_n^n \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} \;=\; \begin{pmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_n) \end{pmatrix}$$

The matrix V is known as a Vandermonde matrix.

Vandermonde Matrices with Monomial Basis

For the monomial basis the condition number of the Vandermonde matrix V grows exponentially with the degree of the polynomial which may lead to inaccuracies when solving the linear system. In addition, V is dense so it requires $\mathcal{O}(n^3)$ operations to solve the linear system.

[Recall that if $Ax = b$ then the perturbed system $(A + \delta A)(x + \delta x) = b + \delta b$ has a solution for which

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{1}{1-\epsilon}\left(\epsilon + \kappa(A)\frac{\|\delta b\|}{\|b\|}\right)$$

where $\epsilon := \|A^{-1}\delta A\| \leq 1$ and $\kappa(A) = \|A\|/\|A^{-1}\|$ is the condition number of the matrix A.]

Vandermonde Matrices with General Basis

We may write $p_n(x)$ in a more general form as

$$p_n(x) = a_0 c_0(x) + a_1 c_1(x) + \ldots + a_n c_n(x) \, ,$$

where the functions $c_i(x)$ form a basis for the space of polynomials of degree n. The interpolation conditions can again be written as a Vandermonde system:

$$\begin{pmatrix} c_0(x_0) & c_1(x_0) & \cdots & c_n(x_0) \\ c_0(x_1) & c_1(x_1) & \cdots & c_n(x_1) \\ \vdots & \vdots & & \vdots \\ c_0(x_n) & c_1(x_n) & \cdots & c_n(x_n) \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_n) \end{pmatrix}$$

For general basis functions V is still a dense matrix but if we choose $c_k(x) = L_{n,k}(x)$, known as the *Lagrange basis*, then V reduces to the identity matrix.

Evaluation of Interpolant with Lagrange Basis

In the Lagrange basis we have

$$p_n(x) = \sum_{k=0}^{n} f(x_k) L_{n,k}(x)$$

where

$$L_{n,k}(x) = \prod_{\substack{j=0 \\ j\neq k}}^{n} \frac{x - x_j}{x_k - x_j} .$$

Again, a naïve evaluation takes $\mathcal{O}(n^2)$ evaluations. However, we can do better.

Evaluation of Interpolant with Lagrange Basis (cont'd)

Define

$$\Pi(x) \;=\; \prod_{k=0}^{n}(x - x_k)\;.$$

Then we have

$$L_{n,k}(x) \;=\; \frac{\Pi(x)}{x - x_k}\omega_k \tag{2}$$

where

$$\omega_k \;=\; \frac{1}{\prod_{j\neq k}(x_k - x_j)} \;=\; \frac{1}{\Pi'(x_k)}\;.$$

Note that the ω_k are independent of x and the function to be interpolated — they just depend on the nodes. We may then write

$$p_n(x) \;=\; \Pi(x)\sum_{k=0}^{n}\frac{\omega_k}{x - x_k}f(x_k)\;.$$

This is the *first barycentric interpolation formula* and the values ω_k are known as the barycentric weights.

Second Barycentric Interpolation Formula

Since the Lagrange interpolant is unique, we know that we can write the interpolant of 1 as

$$\sum_{k=0}^{n} L_{n,k}(x) \;=\; 1\,.$$

Then, by Equation (2), we have

$$\sum_{k=0}^{n} L_{n,k}(x) \;=\; \Pi(x)\sum_{k=0}^{n}\frac{\omega_k}{x-x_k} \;=\; 1\,,$$

so that we may write

$$\Pi(x) \;=\; \left(\sum_{k=0}^{n}\frac{\omega_k}{x-x_k}\right)^{-1}\,.$$

This allows us to eliminate $\Pi(x)$ in the first barycentric formula to get the second barycentric interpolation formula

$$p_n(x) \;=\; \frac{\sum_{k=0}^{n}\frac{\omega_k}{x-x_k}f(x_k)}{\sum_{k=0}^{n}\frac{\omega_k}{x-x_k}}\,.$$

Second Barycentric Interpolation Formula with Chebyshev Grids

The second barycentric interpolation formula has the advantage that if the ω_k all contain a large factor this can be cancelled in the expression for p_n. For example, for the Chebyshev nodes it is known that

$$\omega_k = (-1)^k \frac{2^{n-1}}{n}$$

for $k = 1, \ldots, n-1$ and half of this for $k = 0$ and $k = n$. Clearly the factor $2^{n-1}/n$ can get very large, but this term cancels out so we have

$$p_n(x) = \sum_{k=0}^{n}{}' \frac{(-1)^k f(x_k)}{x - x_k} \Big/ \sum_{k=0}^{n}{}' \frac{(-1)^k}{x - x_k} .$$

Here $\sum'$ means the first and last terms are multiplied by $1/2$. (Note there are issues at the nodes but the value of the interpolant there is known.)

Summary

For the Lagrange interpolation problem:

- the monomial basis is a bad idea as it leads to ill-conditioned Vandermonde matrices;
- equally spaced points can be a bad idea due to the Runge phenomenon;
- the Chebyshev points work well for functions with some smoothness;
- the Lagrange basis works well as the Vandermonde matrix is the identity matrix;
- the second barycentric formula provides an efficient way to evaluate the Lagrange interpolant based on Chebyshev nodes.