Interpolation

Big Idea. An interpolating function provides information about values between points and beyond the range of data. There are infinitely many different ways to interpolate a set of data! Polynomial interpolation is the simplest method whereas cubic spline interpolation provides much more flexibility

Interpolating Functions

An interpolating function (or interpolant) for points $(t_0,y_0),\ldots,(t_d,y_d)$ is a function f(t) such that $f(t_k)=y_k$ for $k=0,\ldots,d$

Note that there are *infinitely* many different interpolating functions for any set of data. We impose different kinds of constraints on the form of the interpolant to ensure there is a unique solution which is meaningful for the data and can be computed efficiently. In the following sections we describe two kinds of interpolation *polynomial* and *cubic spline*

Polynomial interpolation

A polynomial of degree (at most) d is a function of the form

$$p(t) = c_0 + c_1 t + \dots + c_d t^d$$

where $c_0, c_1, \dots, c_d \in \mathbb{R}$ the collection of all polynomials of degree (at most) d is denoted by

$$\mathbb{P}_d = \{c_0+c_1t+, \cdots + c_dt^d: c_0, c_1, \ldots, c_d \in \mathbb{R}\}$$

Note that \mathbb{P}_d is a vector space of dimension d+1

Fix a positive integer d and let $\phi_k(t) = t^k$ for $k = 0, \dots, d$ in other words consider the polynomials

$$\phi_0(t)=1,\quad \phi_{1(t)=t},\quad \dots,\quad \phi_d(t)=t^d$$

Since a polynomial p(t) of degree (at most) d is of the form

$$p(t) = c_0 \phi_0(t) + c_1 \phi_1(t) + \cdots + c_d \phi_d(t)$$

we see that the polynomials $\phi_k(t)$ form a basis of \mathbb{P}_d . Then the set $\{\phi_0(t), \dots, \phi_d(t)\}$ is called the **monomial basis** of \mathbb{P}_d

Given d+1 points $(t_0, y_0), \ldots, (t_d, y_d)$, polynomial interpolation (with respect to the monomial basis) is a polynomial of the form

$$p(t) - c_0 + c_1 t + \dots + t_d t^d$$

such that $p(t_k) = y_k$ for each $k = 0, \dots d$. Each point defines an equation

$$p(t_k) = y_k \quad , \quad k = 0, \ldots, d$$

therefore the coefficients c_0, c_1, \dots, c_d satisfy the system of linear equations $A \vec{c} = \vec{y}$ where

$$A = egin{bmatrix} 1 & t_0 & \dots & t_0^d \ 1 & t_1 & \dots & t_1^d \ dots & dots & \ddots & dots \ 1 & t_d & \dots & t_d^d \end{bmatrix} \quad c = egin{bmatrix} c_0 \ c_1 \ dots \ c_d \end{bmatrix} \quad y = egin{bmatrix} y_0 \ y_1 \ dots \ y_d \end{bmatrix}$$

The matrix A is called the Vandermonde matrix for t_0, \ldots, t_d .



Let A be the Vandermonde matrix for points $(t_0, y_0), \ldots, (t_d, y)d$ Then

$$\det(A) = \prod_{0 \le i < j \le d} (t_j - t_i)$$

= EXAMPLE

Compute the formula for the cases d=1 and d=2

d = 1

$$\det(\begin{bmatrix}1 & t_0 \\ 1 & t_1\end{bmatrix}) = t_1 - t_0$$

d = 2

Use $\det(A) = \det(U)$ from the LU decomposition. Compute

$$\begin{bmatrix} 1 & t_0 & t_0^2 \\ 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & t_0 & t_0^2 \\ 0 & t_1 - t_0 & t_1^2 - t_0^2 \\ 0 & t_2 - t_0 & t_2^2 - t_0^2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & t_0 & t_0^2 \\ 0 & t_1 - t_0 & t_1^2 - t_0^2 \\ 0 & 0 & \star \end{bmatrix}$$

where
$$\star=t_2^2-t_0^2-rac{t_2-t_0}{t_1-t_0}(t_1^2-t_0^2)=(t_2-t_0)(t_2-t_1)$$

The determinant is the product of the diagonal entries of U $\det(U)=(t_1-t_0)(t_2-t_0)(t_2-t_1)$

Consider d+1 data points $(t_0,y_0),\ldots,(t_dy_d)$ such that $t_i\neq j$ for $i\neq j$. There exist a unique polynomial p(t) of degree (at most) d such that $p(t_k)=y_k$ for each k



The formula for the determinant of the Vandermonde matrix shows that $\det(A) \neq_9$ when the values t_k are distinct therefore it is invertible and so there is a unique solution of the system $A\vec{c} = \vec{y}$

= EXAMPLE

Find the unique polynomial of degree 3 which interpolates the points

$$(0,-1),(1,-1),(2,1),(3,-1)$$

Solve the system $A\vec{c} = \vec{y}$

$$A = egin{bmatrix} 1 & 0 & 0 & 0 \ 1 & 1 & 1 & 1 \ 1 & 2 & 4 & 8 \ 1 & 3 & 9 & 27 \end{bmatrix} \quad y = egin{bmatrix} -1 \ -1 \ 1 \ -1 \end{bmatrix}$$

Lets compute the LU decomposition of the Vandermonde matrix \boldsymbol{A} so that we can reuse the result in our other examples

$$L = egin{bmatrix} 1 & 0 & 0 & 0 \ 1 & 1 & 0 & 0 \ 1 & 2 & 1 & 0 \ 1 & 3 & 3 & 1 \end{bmatrix} \quad U = egin{bmatrix} 1 & 0 & 0 & 0 \ 0 & 1 & 1 & 1 \ 0 & 0 & 2 & 6 \ 0 & 0 & 0 & 6 \end{bmatrix}$$

Solve the systems Lx = y and Uc = x to find

$$c = \begin{bmatrix} -1 & -3 & 4 & -1 \end{bmatrix}$$

Therefore the interpolating polynomial is $p(t) = -1 - 3t + 4t^2 - t^3$

Note. The condition number of a Vandermonde matrix gets very large as the size of the matrix increases.

This means that interpolation by the monomial basis is very sensitive to change on the data for polynomials of large degree. For example, for 11 equally spaced points $t_0=0,\ldots,t_{10}=10$ the Vandermonde matrix A is 11×11 and has a condition number larger than 10^{12}

Cubic spline interpolation

Consider N+1 points $(t_0,y_0),\ldots,(t_N,y_N)$. A cubic spline is a function p(t) defined piecewise by N cubic polynomials $p_1(t),\ldots,p_N(t)$ where

$$p_k(t) = a_k(t-t_{k-1})^3 + b_k(t-t_{k-1})^2 + c_k(t-t_{k-1}) + d_k \quad , \quad t \in [t_{l-1},t_k]$$

such that p(t), p'(t), p''(t) are continuous functions

Each polynomial $p_k(t)$ is defined by our four coefficients therefore we require 4N equations to specify the 4N unknowns

Interpolation at the left endpoints yields N equations:

$$p_k(t_{k-1})=y_{k-1}$$
 , $k=1,\ldots,N$

Interpolation at the right endpoints yields N equations:

$$p_k(t_k) = y_k \quad , \quad k = 1, \dots N$$

Continuity of p'(t) yields N-1 equations:

$$p_k'(t_k) = p_{k+1}'(t_k) \quad , \quad k = 1, \dots, N-1$$

Continuity of p''(t) yields N-1 equations:

$$p_k''(t_k) = p_{k+1}''(t_k) \quad , \quad k = 1, \dots, N-1$$

The conditions impose only 4N-2 equations therefore we need two more to determine the cubic spline uniquely. There are different choices such as the natural cubic spline and the "not-a-knot" condition

DEFINTION

A natural cubic spline satisfies $p_1^{\prime\prime}(t_0)=p_N^{\prime\prime}(t_N)=0$

Represent a cubic spline p(t) by the coefficient matrix

$$C = egin{bmatrix} a_1 & a_2 & \dots & a_N \ b_1 & b_2 & \dots & b_N \ c_1 & c_2 & \dots & c_N \ d_1 & d_2 & \dots & d_N \end{bmatrix}$$

where the $k{
m th}$ column consists of the coefficients for the $k{
m th}$ cubic polynomial in the spline

Consider N+1 points $(t_0,y_0),\ldots,(t_N,y_N)$ with $t_i\neq t_j$ for $i\neq j$. The unique natural cubic spline p(t) which interpolates the points is given by the matrix C

$$C = egin{bmatrix} a_1 & a_2 & \dots & a_N \ b_1 & b_2 & \dots & b_N \ c_1 & c_2 & \dots & c_N \ d_1 & d_2 & \dots & d_N \end{bmatrix}$$

where $d_k=y_{k-1}$ for $k=1,\ldots,N$ and the coefficients $a_1,b_1,c_1,\ldots,a_N,b_N,c_N$ are the solutions of the linear system

					$\mid a_1 \mid$	
$A(L_1)$	B				b_1	
	$A(L_2)$	B			c_1	
		٠	٠		:	
			$A(L_{N-1})$	B	$ a_N $	
\overline{T}				$\overline{\hspace{1em}V\hspace{1em}}$	b_N	

$$egin{bmatrix} a_1 \ b_1 \ c_1 \ dots \ a_N \ b_N \ c_N \end{bmatrix} = egin{bmatrix} y_1 - y_0 \ 0 \ 0 \ dots \ y_N - y_{N-1} \ 0 \ 0 \end{bmatrix}$$

where $L_k = t_k - t_{k-1}$ is the length of the subinterval $[t_{k-1}, t_k]$ and

$$A(L) = egin{bmatrix} L^3 & L^2 & L \ 3L^2 & 2L & 1 \ 6L & 2 & 0 \end{bmatrix} \quad B = egin{bmatrix} 0 & 0 & 0 \ 0 & 0 & -1 \ 0 & -2 & 0 \end{bmatrix}$$

$$T = egin{bmatrix} 0 & 0 & 0 \ 0 & 2 & 0 \ 0 & 0 & 0 \end{bmatrix} \quad V = egin{bmatrix} L_N^3 & L_N^2 & L_N \ 0 & 0 & 0 \ 6L_N & 2 & 0 \end{bmatrix}$$

EXAMPLE

Let p(t) be the natural cubic spline which interpolates the data

$$(0,1),(1,3),(2,8),(3,10),(4,9),(5,-1),(6,-17)$$

Suppose the coefficient matrix of p(t) is

$$C = egin{bmatrix} 1 & -2 & 1 & a_4 & 1 & 1 \ 0 & 3 & -3 & b_4 & -6 & -3 \ 1 & 4 & 4 & c_4 & -5 & -14 \ 1 & 3 & 8 & 10 & 9 & -1 \end{bmatrix}$$

Determine the coefficients a_4, b_4, c_4 and then compute the value p''(2.5)First note that the t values are equally spaced therefore $L_k = t_k - t_{k-1} = 1$ for each k. Continuity of p(t), p't and p''(t) yields the equations Interpolation

$$egin{aligned} p_k(t_k) &= p_{k+1}(t_k) \Longrightarrow \ a_k + b_k + c_k + d_k = d_{k+1} \ p_k'(t_k = p_{k+1}''(t_k) \Longrightarrow \ 3a_k + 2b_k + c_k = c_{k+1} \ p_k''(t_k) &= p_{k+1}''(t_k) \Longrightarrow \ 6a_k + 2b_k = 2b_{k+1} \end{aligned}$$

From these equations we can read that $b_4=0, \quad c_4=0, \quad a_4=-2$ from the equation $6a_4+2b_4=2b_5$

The value t=2.5 is in the interval $[t_2,t_3]$ therefore we compute $p''(2.5)=p_3''(2.5)=6a_3(2.5-2)+2b_3=-3$

Note the condition number of the matrix for constructing the natural cubic spline does not increase as drastically with the number of points N+1 as compared with the Vandermonde matrix. For example, for 11 equally spaced points $t_0=0,\ldots,t_{10}=10$, the Vandermonde matrix is 11 by 11 and has the $\operatorname{cond}(A)\approx 10^{12}$ whereas the cubic spline matrix is 30 by 30 and the condition number is only around 33.