Newton Polynomial

https://wolog.xyz

1. Divided Differences

1.1. Introduction

For any real function g and any nonempty set S of abscissae, we know there uniquely exists a polynomial of degree at most $|S|-1$ that agrees with g at all $a \in S$, and in this article we denote this polynomial as $P^g_S(x)$. $P^g_\varnothing(x)$ is defined to be the zero polynomial for any function g.

Given a real function f, a finite set A of abscissae and an additional abscissa b, suppose $P^f_A(x)$ is known and we want to find $P^f_{A\cup\{b\}}(x)$. To make use of $P^f_A(x)$, observe that $P^f_{A\cup\{b\}}(x) - P^f_A(x)$ is a polynomial of degree at most $|A|$ that vanishes at every $a \in A$, so it must be equal to $C \prod_{a\in A}(x-a)$ for some constant C, whose uniqueness follo ws the uniqueness of the polynomial. That is, there uniquely exists a constant C such that

$$P^f_{A\cup\{b\}}(x) = P^f_A(x) + C \prod_{a\in A}(x-a)$$

What remains is to study this constant.

When $A = \{x_0, x_1, \cdots, x_{n-1}\}$ and $b = x_n$, this constant is denoted as $f[x_0, x_1, \cdots, x_n]$, and called a divided difference for a reason that will be shown later. So here comes our definition for divided differences:

- The divided difference $f[x_0, x_1, \cdots, x_n]$ is the unique number satisfying

$$P^f_{\{x_0,x_1,\cdots,x_n\}}(x) = P^f_{\{x_0,x_1,\cdots,x_{n-1}\}}(x) + f[x_0, x_1, \cdots, x_n](x-x_0)(x-x_1)\cdots(x-x_{n-1})$$

Observe that, on the right-hand side of the equation in this definition, $P^f_{\{x_0,x_1,\cdots,x_{n-1}\}}(x)$'s degree is less than n, so the factor x^n can only come from expanding $(x-x_0)(x-x_1)\cdots(x-x_{n-1})$. This brings us the following characterization:

- $f[x_0, x_1, \cdots, x_n]$ is the coefficient of x^n in $P^f_{\{x_0,x_1,\cdots,x_n\}}(x)$

This characterization shows the symmetry of divided differences: They don't depend on the order of abscissae.

1.2. A Straightforward & Symmetric Formula

By the characterization of divided differences that we just established, a formula for the coefficient of x^n in $P^f_{\{x_0,x_1,\cdots,x_n\}}(x)$ will be a formula for $f[x_0, x_1, \cdots, x_n]$. We can use the Lagrange form of the interpolation polynomial, usually called the Lagrange polynomial, to find it.

For a recap, building blocks of the Lagrange polynomial are

$$l_i(x) = \prod_{\substack{0\le k\le n \\ k\ne i}} \frac{x-x_k}{x_i-x_k}, \quad i = 0, 1, \cdots, n$$

So each $l_i(x)$ is a polynomial of degree n that, by design, attains 1 at x_i but vanishes at all the other abscissae. The Lagrange polynomial is then constructed as

$$\sum_{i=0}^{n} f(x_i) l_i(x)$$

which clearly coincides with $f(x)$ at e very abscissa.

Our focus is on the coefficient of x^n. x^n has the following coefficient in $l_i(x)$:

$$\frac{1}{\prod\limits_{\substack{0\le k\le n\\ k\ne i}}(x_i - x_k)}$$

So in total, the coefficient of x^n is:

$$f[x_0, x_1, \cdots, x_n] = \sum_{i=0}^{n} \frac{f(x_i)}{\prod\limits_{\substack{0\le k\le n\\ k\ne i}}(x_i - x_k)}$$

1.3. The "Divided Difference" Recurrence Relation

We've seen that a polynomial can be extended to pass through an additional data point by adding a term involving a divided difference. One interesting thing is that adding multiple data points in different orders will produce apparently different polynomials. Suppose the abscissae are x_0, x_1, ..., x_n ($n > 0$) and $P^f_{\{x_1,x_2,\cdots,x_{n-1}\}}(x)$ is known. If we add $(x_0, f(x_0))$ first and $(x_n, f(x_n))$ next, we have

$$\begin{aligned}P^f_{\{x_0,x_1,\cdots,x_n\}}(x) &= P^f_{\{x_1,x_2,\cdots,x_{n-1}\}}(x) + f[x_0, x_1, \cdots, x_{n-1}](x - x_1)(x - x_2)\cdots(x - x_{n-1}) \\ &\qquad + f[x_0, x_1, \cdots, x_n](x - x_0)(x - x_1)\cdots(x - x_{n-1}) \\ &= P^f_{\{x_1,x_2,\cdots,x_{n-1}\}}(x) + \big(f[x_0, x_1, \cdots, x_{n-1}] + f[x_0, x_1, \cdots, x_n](x - x_0)\big)(x - x_1)(x - x_2)\cdots(x - x_{n-1})\end{aligned}$$

But if we add $(x_n, f(x_n))$ first and $(x_0, f(x_0))$ next, we have

$$\begin{aligned}P^f_{\{x_0,x_1,\cdots,x_n\}}(x) &= P^f_{\{x_1,x_2,\cdots,x_{n-1}\}}(x) + f[x_1, x_2, \cdots, x_n](x - x_1)(x - x_2)\cdots(x - x_{n-1}) \\ &\qquad + f[x_0, x_1, \cdots, x_n](x - x_1)(x - x_2)\cdots(x - x_{n-1})(x - x_n) \\ &= P^f_{\{x_1,x_2,\cdots,x_{n-1}\}}(x) + \big(f[x_1, x_2, \cdots, x_n] + f[x_0, x_1, \cdots, x_n](x - x_n)\big)(x - x_1)(x - x_2)\cdots(x - x_{n-1})\end{aligned}$$

Comparing them, we obtain the following equation:

$$f[x_0, x_1, \cdots, x_{n-1}] + f[x_0, x_1, \cdots, x_n](x - x_0) = f[x_1, x_2, \cdots, x_n] + f[x_0, x_1, \cdots, x_n](x - x_n)$$

Therefore,

$$f[x_0, x_1, \cdots, x_n] = \frac{f[x_1, x_2, \cdots, x_n] - f[x_0, x_1, \cdots, x_{n-1}]}{x_n - x_0}$$

The initial values are $f[x_i] = f(x_i)$ for $i = 0, 1, \cdots, n$. This recurrence relation is what divided differences are named for.[1]

2. The Newton Polynomial

Knowing how divided differences can be used to extend an interpolation polynomial, now let's repeat the process to construct an interpolation polynomial from the ground up. Begin with the zero polynomial and add a term for each of $(x_0, f(x_0))$, $(x_1, f(x_1))$, ..., $(x_n, f(x_n))$, we end up with

$$f[x_0] + f[x_0, x_1](x - x_0) + \cdots + f[x_0, x_1, \cdots, x_n](x - x_0)(x - x_1)\cdots(x - x_{n-1})$$

This is the Newton form of the interpolation polynomial, conventionally referred to as the Newton polynomial.

The recurrence relation for divided differences provides an efficient method of computing all coefficients in a Newton polynomial, as demonstrated in the following tabular form:

---

[1] It is in fact the prevalent definition of divided differences. But I think using it as the definition obfuscates the nature of divided differences (the same can even be said of using "divided difference" as the name), and typically leads to unintuitive derivations of the Newton polynomial like the one in the Wikipedia article *Newton Polynomial*.

$$\begin{array}{c|cccc}
x_0 & f[x_0] & & & \\
 & & \frac{f[x_1]-f[x_0]}{x_1-x_0} = f[x_0, x_1] & & \\
x_1 & f[x_1] & & \frac{f[x_1,x_2]-f[x_0,x_1]}{x_2-x_0} = f[x_0, x_1, x_2] & \\
 & & \frac{f[x_2]-f[x_1]}{x_2-x_1} = f[x_1, x_2] & & \ldots \\
x_2 & f[x_2] & & \ldots & \\
 & & \ldots & & \\
\ldots & \ldots & & &
\end{array}$$

Besides the efficiency of construction, the Newton polynomial is efficient to evaluate, as $x - x_0$, $(x - x_0)(x - x_1)$, ..., $(x - x_0)(x - x_1)\cdots(x - x_{n-1})$ can be evaluated successively.

3. An Unusual Application

Interpolation polynomials are commonly thought of as a numerical tool for approximation. But when the function that's sampled for data points is itself a polynomial, an interpolation polynomial can match the function exactly provided enough data points. Below I will show how this can be useful mathematically.

Let's consider the sum

$$S_k(n) = 1^k + 2^k + \cdots + n^k \quad (k \in \mathbf{N})$$

Here are some well known cases:

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

$$1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

$$1^3 + 2^3 + \cdots + n^3 = \frac{n^2(n+1)^2}{4}$$

We're going to find a general formula in a novel way below.

Let $p(x)$ be the polynomial of degree not greater than $k+1$ that agrees with $S_k(x)$ at the $k+2$ abscissae $0, 1, \ldots, k+1$. Consider the polynomial $q(x) = p(x) - p(x-1)$, which satisfies

- $q(n) = p(n) - p(n-1) = n^k$ for $n = 1, 2, \cdots, k+1$.

In the polynomial subtraction $p(x) - p(x-1)$, x^{k+1} from $p(x)$ is canceled by expanding $(x-1)^{k+1}$ from $p(x-1)$, so

- $q(x)$'s degree is not greater than k.

Therefore, $q(x)$ must be the polynomial x^k, so

- $p(n) - p(n-1) = q(n) = n^k$ for $n \in \mathbf{N} \setminus \{0\}$.

It follows that $p(n) = 1^k + 2^k + \cdots + n^k$ for $n \in \mathbf{N}$. So $S_k(x)$ can be equated with an interpolation polynomial of degree not greater than $k+1$. By using the Newton polynomial to interpolate, we get the following formula:

$$S_k(n) = \sum_{i=0}^{k+1} S_k[0, 1, \cdots, i]\, n(n-1)\cdots(n-i+1)$$

The Lagrange polynomial will work as well, but with the Newton polynomial we can make use of the fact that $S_k[n-1, n] = n^k$ for $n = 1, 2, \cdots, k+1$ to a void evaluating S_k.