4. Polynomial Approximation

In many problems occuring in applications, we are confronted with the problem of approximating an unknown function from various types of information we are given about the function. A simple example occurs when we have measurements of the function (possibly by an experiment) at discrete time intervals and want to have an approximation of the function at times for which we have no measurements. A more complicated situation occurs when we know that the function we seek satisfies a certain differential equation, but the equation is not one where an analytical technique can be applied to find its solution. In these cases, we may seek an approximation to the function by a simple function that can be easily evaluated on a computer. One reason for the use of polynomials as approximating functions is that they are easy to evaluate on a computer. Another important reason is that they have good approximation properties, made more precise in the following theorem.

4.1. **Weierstrass Approximation Theorem.** If $f(x)$ is continuous on a finite interval $[a, b]$, then given $\epsilon > 0$, there exists n depending on ϵ and a polynomial $P_n(x)$ of degree $\leq n$ such that $|f(x) - P_n(x)| \leq \epsilon$ for all $x \in [a, b]$.

The proof uses Bernstein polynomials: These polynomials are defined on the interval $[0, 1]$ by: $B_n(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f(k/n)$. One can show that $\lim_{n\to\infty} B_n(x) = f(x)$ uniformly in $[0, 1]$. If $y \in [a, b]$, we can reduce the problem to the interval $[0, 1]$ by the change of variable $x = (y - a)/(b - a)$. The proof is not constructive in the sense that we do not know how large n has to be to achieve a given accuracy ϵ. However, the theorem does tell us we can approximate continuous functions to any accuracy using polynomials.

4.2. **Forms of Polynomials.** Probably the most familiar way to write a general polynomial of degree $\leq n$ is to use the form

$$P_n(x) = a_0 + a_1 x + \cdots + a_n x^n,$$

where the a_i are arbitary constants. In such a formula, we can think of the a_i as the *degrees of freedom* of the polynomial, since these the terms we are free to choose in the formula. In fact, the a_i are related to the value of P_n and its deratives at $x = 0$, i.e.,

$$a_0 = P_n(0), \qquad a_1 = P_n'(0), \qquad a_k = P_n^{(k)}(0)/k!.$$

Thus, in this formula, the degrees of freedom of the polynomial are $P_n^{(k)}(0)/k!$.

This representation of a polynomial is the one used when we write down the first $n + 1$ terms of the Taylor series expansion of a function f, i.e.,

$$T_n(x) = f(0) + f'(0)x + \cdots + \frac{f^{(n)}(0)}{n!} x^n.$$

Thus, the Taylor polynomial $T_n(x)$ is the polynomial $P_n(x)$ of degree $\leq n$ for which

$$P_n(0) = f(0), \quad P_n'(0) = f'(0), \quad \cdots, \quad P_n^{(n)}(0) = f^{(n)}(0).$$

If we consider the Taylor series of f about the point $x = c$, then we are using the polynomial representation:

$$P_n(x) = b_0 + b_1(x - c) + \cdots + b_n(x - c)^n.$$

In this case, the degrees of freedom b_k correspond to constant multiples of the value and derivatives of P_n at the point $x = c$, i.e.,

$$b_0 = P_n(c), \quad b_1 = P_n'(c), \quad \cdots, \quad b_n = P_n^{(n)}(c)/n!,$$

where

$$P_n(c) = f(c), \quad P_n'(c) = f'(c), \quad \cdots, \quad P_n^{(n)}(c) = f^{(n)}(c).$$

While these two representations of a polynomial of degree $\leq n$ are useful in some application, there are other representations that are more useful in other applications. For example, a classical problem in data fitting is the following:

4.3. **Polynomial interpolation:** Given $n+1$ distinct points $x_0, \cdots, x_n$ and function values $f(x_0), \cdots, f(x_n)$, find a polynomial $P_n(x)$ of degree $\leq n$ satisfying $P_n(x_j) = f(x_j)$, $j = 0, 1, \ldots, n$. We say P_n interpolates f at $x_0, \ldots, x_n$. If we use the representation $P_n(x) = a_0 + a_1 x + \cdots + a_n x^n$, then to determine the a_i, we would need to solve the following linear system of $n + 1$ equations:

$$\begin{aligned} a_0 + a_1 x_0 + \cdots + a_n x_0^n &= f(x_0) \\ a_0 + a_1 x_1 + \cdots + a_n x_1^n &= f(x_1) \\ &\cdots \\ a_0 + a_1 x_n + \cdots + a_n x_n^n &= f(x_n) \end{aligned}$$

for the $n + 1$ unknowns $a_0, \cdots, a_n$. Instead, if we start from another representation of a polynomial, called the *Lagrange form*, then we can write down the solution immediately.

4.4. **Lagrange form of the interpolating polynomial.** : Define for $k = 0, 1, \ldots, n$:

$$L_{k,n}(x) = \prod_{\substack{j=0 \\ j \neq k}}^{n} \frac{(x - x_j)}{(x_k - x_j)}, \quad n \geq 1, \qquad L_{0,0}(x) = 1.$$

Claim: $P(x) = \sum_{k=0}^{n} L_{k,n}(x) f(x_k)$ is a polynomial of degree $\leq n$ that interpolates f at $x_0, \ldots, x_n$. This representation is called the Lagrange form of the interpolating polynomial.

Another way of thinking about this form of a polynomial is that we are writing $P(x) = \sum_{k=0}^{n} L_{k,n}(x) P(x_k)$. The functions $L_{k,n}(x)$ are fixed and the polynomial $P(x)$ is then uniquely determined by its values $P(x_k)$ at the $n+1$ points $x_0, \ldots, x_n$. These values $P(x_0), \ldots, P(x_n)$ are the degrees of freedom of $P(x)$ (in this form of the polynomial), i.e., they are the quantities we are free to choose that uniquely determine $P(x)$. Once we know we can write any polynomial in this form, it is then easy to solve the interpolation problem, i.e., we simply replace $P(x_k)$ by $f(x_k)$.

Observe that $L_{k,n}(x)$ is a product of n monomials in x and hence is a polynomial of degree $\leq n$. Taking linear combinations still gives a polynomial of degree $\leq n$.

Next: $L_{k,n}(x_i) = 0, i \neq k$. $L_{k,n}(x_i) = 1, i = k$. Hence, $P(x_i) = f(x_i)$.

Uniqueness: Suppose $Q(x)$ is another interpolating polynomial of degree $\leq n$. Then $R(x) = P(x) - Q(x)$ is of degree $\leq n$ and equals zero at $x_0, \ldots, x_n$. Since $R(x) = 0$ at the distinct points $x_1, \ldots, x_n$, $R(x)$ has the form $R(x) = A(x - x_1) \cdots (x - x_n)$ for some constant A. Then $R(x_0) = 0$ implies $A = 0$, so $R(x) = 0$.

Note: There exists other polynomials of degree $d > n$ interpolating f at $x_0, \ldots, x_n$.

Examples: $n = 0$: $P_0(x) = f(x_0)$. $\qquad n = 1$: $P_1(x) = \frac{x-x_1}{x_0-x_1} f(x_0) + \frac{x-x_0}{x_1-x_0} f(x_1)$.

4.5. **Newton form of the interpolating polynomial.** One problem with the Lagrange form of the interpolating polynomial is that if we have already computed P_{n-1} and now add one additional interpolation point, we have to recompute everthing. Thus, we consider a second form of the interpolating polynomial, known as the Newton form.

Note we can also write:

$$P_1(x) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0).$$

This is an example of the Newton form of the interpolating polynomial. In general, we want to write $P_n(x) = P_{n-1}(x) + Q_n(x)$, where P_{n-1} interpolates f at $x_0, \ldots, x_{n-1}$ and $Q_n(x)$ has a simple form.

By the definitions of P_n and P_{n-1}, $Q_n(x)$ is of degree $\leq n$ and $= 0$ at $x_0, \ldots, x_{n-1}$. Hence, $Q_n(x) = A_n(x - x_0) \cdots (x - x_{n-1})$ for some constant A_n. Hence, we need only determine A_n.

Now $f(x_n) = P_n(x_n) = P_{n-1}(x_n) + Q_n(x_n)$. Hence,

$$\begin{aligned}
A_n &= \frac{f(x_n) - P_{n-1}(x_n)}{\prod_{j=0}^{n-1}(x_n - x_j)} \\
&= f(x_n) - \sum_{k=0}^{n-1} \left[\prod_{\substack{j=0 \\ j \neq k}}^{n-1} \frac{x_n - x_j}{x_k - x_j} \right] f(x_k) / \prod_{j=0}^{n-1}(x_n - x_j) \\
&= \frac{f(x_n)}{\prod_{j=0}^{n-1}(x_n - x_j)} - \sum_{k=0}^{n-1} \frac{f(x_k)}{\prod_{\substack{j=0 \\ j \neq k}}^{n-1}(x_k - x_j)} \cdot \frac{1}{x_n - x_k} \\
&= \sum_{k=0}^{n} \frac{f(x_k)}{\prod_{\substack{j=0 \\ j \neq k}}^{n}(x_k - x_j)}. \qquad (4.1)
\end{aligned}$$

We refer to A_n given by this formula as the n^{th} divided difference of f with respect to $x_0, \ldots, x_n$ and denote it by $f[x_0, x_1, \ldots, x_n]$.