Suppose there are two different polynomials of degree $n$ that agree at $n + 1$ distinct points. Call these $P_n(x)$ and $Q_n(x)$, and write their difference:

$$D(x) = P_n(x) - Q_n(x),$$

where $D(x)$ is a polynomial of at most degree $n$. But because $P$ and $Q$ match at the $n + 1$ points, their difference $D(x)$ is equal to zero at all $n + 1$ of these $x$-values; that is, $D(x)$ is a polynomial of degree $n$ at most but has $n + 1$ distinct zeros. However, this is impossible unless $D(x)$ is identically zero. Hence $P_n(x)$ and $Q_n(x)$ are not different—they must be the same polynomial.

A most important consequence of this uniqueness property of interpolating polynomials is that their error terms are also identical (though we may want to express the error terms in different forms). We only have to derive the error term for one form of interpolating polynomial to have the error term for all forms of interpolating polynomials.

## Error of Interpolation

When we fit a polynomial $P_n(x)$ to some data points, it will pass exactly through those points, but between those points $P_n(x)$ will not be precisely the same as the function $f(x)$ that generated the points (unless the function is that polynomial). How much is $P_n(x)$ different from $f(x)$? How large is the error of $P_n(x)$?

We begin the development of an expression for the error of $P_n(x)$, an $n$th-degree interpolating polynomial, by writing the error function in a form that has the known property that it is zero at the $n + 1$ points, from $x_0$ through $x_n$, where $P_n(x)$ and $f(x)$ are the same. We call this function $E(x)$:

$$E(x) = f(x) - P_n(x) = (x - x_0)(x - x_1) \cdots (x - x_n)g(x).$$

The $n + 1$ linear factors give $E(x)$ the zeros we know it must have, and $g(x)$ accounts for its behavior at values other than at $x_0, x_1, \ldots, x_n$. Obviously, $f(x) - P_n(x) - E(x) = 0$, so

$$f(x) - P_n(x) - (x - x_0)(x - x_1) \cdots (x - x_n)g(x) = 0. \tag{3.2}$$

To determine $g(x)$, we now use the interesting mathematical device of constructing an auxiliary function (the reason for its special form becomes apparent as the development proceeds). We call this auxiliary function $W(t)$, and define it as

$$W(t) = f(t) - P_n(t) - (t - x_0)(t - x_1) \cdots (t - x_n)g(x).$$

Note in particular that $x$ has *not* been replaced by $t$ in the $g(x)$ portion. ($W$ is really a function of both $t$ and $x$, but we are only interested in variations of $t$.) We now examine the zeros of $W(t)$.

Certainly at $t = x_0, x_1, \ldots, x_n$, the $W$ function is zero ($n + 1$ times), but it is also zero if $t = x$ by virtue of Eq. (3.2). There are then a total of $n + 2$ values of $t$ that make $W(t) = 0$. We now impose the necessary requirements on $W(t)$ for the *law of mean value* to hold. $W(t)$ must be continuous and differentiable. If this is so, there is a zero to its derivative $W'(t)$ between each of the $n + 2$ zeros of $W(t)$, a total of $n + 1$ zeros. If $W''(t)$ exists, and we suppose it does, there will be $n$ zeros of $W''(t)$, and likewise $n - 1$ zeros of $W'''(t)$, and so on, until we reach $W^{(n+1)}(t)$, which must have at least one zero in the interval that has $x_0$, $x_n$, or $x$ as endpoints. Call this value of $t = \xi$. We then have

$$W^{(n+1)}(\xi) = 0 = \frac{d^{n+1}}{dt^{n+1}} [f(t) - P_n(t) - (t - x_0) \cdots (t - x_n)g(x)]_{t=\xi} \tag{3.3}$$
$$= f^{(n+1)}(\xi) - 0 - (n+1)!g(x).$$

The right-hand side of Eq. (3.3) occurs because of the following arguments. The $(n + 1)$st derivative $f(t)$, evaluated at $t = \xi$, is obvious. The $(n + 1)$st derivative of $P_n(t)$ is zero because every time any polynomial is differentiated its degree is reduced by one, so that the $n$th derivative is of degree zero (a constant) and its $(n + 1)$st derivative is zero. We apply the same argument to the $(n + 1)$st-degree polynomial in $t$ that occurs in the last term—its $(n + 1)$st derivative is a constant that results from the $t^{n+1}$ term and is $(n + 1)!$. Of course, $g(x)$ is independent of $t$ and goes through the differentiations unchanged. The form of $g(x)$ is now apparent:

$$g(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}, \qquad \xi \text{ between } (x_0, x_n, x).$$

The conditions on $W(t)$ that are required for this development (continuous and differentiable $n + 1$ times) will be met if $f(x)$ has these same properties, because $P_n(x)$ is continuous and differentiable. We now have our error term:

$$E(x) = (x - x_0)(x - x_1) \cdots (x - x_n) \frac{f^{(n+1)}(\xi)}{(n+1)!}, \tag{3.4}$$

with $\xi$ on the smallest interval that contains $\{x, x_0, x_1, \ldots, x_n\}$.

The expression for error given in Eq. (3.4) is interesting but is not always extremely useful. This is because the actual function that generates the $x_i, f_i$ values is often unknown; we obviously then do not know its $(n + 1)$st derivative. We can conclude, however, that if the function is "smooth," a low-degree polynomial should work satisfactorily. (The smaller the higher derivatives of a function, the smoother it is. For example, for a straight line, all derivatives above the first are zero.) On the other hand, a "rough" function can be expected to have larger errors when interpolated. We can also conclude that extrapolation (applying the interpolating polynomial outside the range of $x$-values employed to construct it) will have larger errors than for interpolation. It also follows that the error is smaller if $x$ is centered within the $x_i$, because this makes the product of the $(x - x_i)$ terms smaller.

Here is an algorithm for interpolation with a Lagrangian polynomial of degree $N$.