3.2.1 Interpolation error formula

Assume that we start with a function $y = f(x)$ and take data points from it to build an interpolating polynomial $P(x)$, as we did with $f(x) = \sin x$ in Example 3.7. The **interpolation error** at x is $f(x) - P(x)$, the difference between the original function that provided the data points and the interpolating polynomial, evaluated at x. The interpolation error is the vertical distance between the curves in Figure 3.3. The next theorem gives a formula for the interpolation error that is usually impossible to evaluate exactly, but often can at least lead to an error bound.

THEOREM 3.4 Assume that $P(x)$ is the (degree $n-1$ or less) interpolating polynomial fitting the n points $(x_1, y_1), \ldots, (x_n, y_n)$. The interpolation error is

$$f(x) - P(x) = \frac{(x-x_1)(x-x_2)\cdots(x-x_n)}{n!} f^{(n)}(c), \tag{3.6}$$

where c lies between the smallest and largest of the numbers $x, x_1, \ldots, x_n$. ■

See Section 3.2.2 for a proof of Theorem 3.3. We can use the theorem to assess the accuracy of the `sin` key we built in Example 3.7. Equation (3.6) yields

$$\sin x - P(x) = \frac{(x-0)\left(x-\frac{\pi}{6}\right)\left(x-\frac{\pi}{3}\right)\left(x-\frac{\pi}{2}\right)}{4!} f''''(c),$$

where $0 < c < \pi/2$. The fourth derivative $f''''(c) = \texttt{sin}\, c$ varies from 0 to 1 in this range. At worst, $|\sin c|$ is no more than 1, so we can be assured of an upper bound on interpolation error:

$$|\sin x - P(x)| \le \frac{\left|(x-0)\left(x-\frac{\pi}{6}\right)\left(x-\frac{\pi}{3}\right)\left(x-\frac{\pi}{2}\right)\right|}{24}|1|.$$

At $x = 1$, the worst-case error is

$$|\sin 1 - P(1)| \le \frac{\left|(1-0)\left(1-\frac{\pi}{6}\right)\left(1-\frac{\pi}{3}\right)\left(1-\frac{\pi}{2}\right)\right|}{24}|1| \approx 0.0005348. \tag{3.7}$$

This is an upper bound for the error, since we used a "worst case" bound for the fourth derivative. Note that the actual error at $x = 1$ was .0004, which is within the error bound given by (3.7). We can make some conclusions on the basis of the form of the interpolation error formula. We expect smaller errors when x is closer to the middle of the interval of x_i's than when it is near one of the ends, because there will be more small terms in the product. For example, we compare the preceding error bound to the case $x = 0.2$, which is near the left end of the range of data points. In this case, the error formula is

$$|\sin 0.2 - P(0.2)| \le \frac{\left|(.2-0)\left(.2-\frac{\pi}{6}\right)\left(.2-\frac{\pi}{3}\right)\left(.2-\frac{\pi}{2}\right)\right|}{24}|1| \approx 0.00313,$$

about six times larger. Correspondingly, the actual error is larger, specifically,

$$|\sin 0.2 - P(0.2)| = |0.19867 - 0.20056| = 0.00189.$$

EXAMPLE 3.8 Find an upper bound for the difference at $x = 0.25$ and $x = 0.75$ between $f(x) = e^x$ and the polynomial that interpolates it at the points $-1, -0.5, 0, 0.5, 1$.

Construction of the interpolating polynomial, shown in Figure 3.4, is not necessary to find the bound. The interpolation error formula (3.6) gives

$$f(x) - P_4(x) = \frac{(x+1)\left(x+\frac{1}{2}\right)x\left(x-\frac{1}{2}\right)(x-1)}{5!} f^{(5)}(c),$$

where $-1 < c < 1$. The fifth derivative is $f^{(5)}(c) = e^c$. Since e^x is increasing with x, its maximum is at the right-hand end of the interval, so $|f^{(5)}| \le e^1$ on $[-1, 1]$. For $-1 \le x \le 1$, the error formula becomes

$$|e^x - P_4(x)| \le \frac{(x+1)\left(x+\frac{1}{2}\right)x\left(x-\frac{1}{2}\right)(x-1)}{5!} e.$$

At $x = 0.25$, the interpolation error has the upper bound

$$\begin{aligned} |e^{0.25} - P_4(0.25)| &\le \frac{(1.25)(0.75)(0.25)(-0.25)(-0.75)}{120} e \\ &\approx .000995. \end{aligned}$$

At $x = 0.75$, the interpolation error is potentially larger:

$$\begin{aligned} |e^{0.75} - P_4(0.75)| &\le \frac{(1.75)(1.25)(0.75)(0.25)(0.25)}{120} e \\ &\approx .002323. \end{aligned}$$

Note again that the interpolation error will tend to be smaller close to the center of the interpolation interval. ◀

EXAMPLE 3 **Error Estimate (5) of Linear Interpolation. Damage by Roundoff. Error Principle**

Estimate the error in Example 1 first by (5) directly and then by the Error Principle (Sec. 19.1).

Solution. (A) ***Estimation by*** (5). We have $n = 1, f(t) = \ln t, f'(t) = 1/t, f''(t) = -1/t^2$. Hence

$$\epsilon_1(x) = (x - 9.0)(x - 9.5)\frac{(-1)}{2t^2}, \qquad \text{thus} \qquad \epsilon_1(9.2) = \frac{0.03}{t^2}.$$

$t = 0.9$ gives the maximum $0.03/9^2 = 0.00037$ and $t = 9.5$ gives the minimum $0.03/9.5^2 = 0.00033$, so that we get $0.00033 \leqq \epsilon_1(9.2) \leqq 0.00037$, or better, 0.00038 because $0.3/81 = 0.003703 \cdots$.

But the error 0.0004 in Example 1 disagrees, and we can learn something! Repetition of the computation there with 5D instead of 4D gives

$$\ln 9.2 \approx p_1(9.2) = 0.6 \cdot 2.19722 + 0.4 \cdot 2.25129 = 2.21885$$

with an actual error $\epsilon = 2.21920 - 2.21885 = 0.00035$, which lies nicely near the middle between our two error bounds.

This shows that the discrepancy (0.0004 vs. 0.00035) was caused by rounding, which is not taken into account in (5).

(B) ***Estimation by the Error Principle.*** We calculate $p_1(9.2) = 2.21885$ as before and then $p_2(9.2)$ as in Example 2 but with 5D, obtaining

$$p_2(9.2) = 0.54 \cdot 2.19722 + 0.48 \cdot 2.25129 - 0.02 \cdot 2.39790 = 2.21916.$$

The difference $p_2(9.2) - p_1(9.2) = 0.00031$ is the approximate error of $p_1(9.2)$ that we wanted to obtain; this is an approximation of the actual error 0.00035 given above.