0.1 Newton's method and the Mean Value Theorem

Newton's method for computing the zeros of functions is a good example of the practical application of the Mean Value Theorem. Let $f(x)$ be a real-valued function on the real line that has two continuous derivatives. We are looking for a *root* of f, i.e., a point $\hat{x}$ such that $f(\hat{x}) = 0$. In Newton's method, which is geometrical, we consider the curve $y = f(x)$. Then the curve crosses the x-axis at the point $(\hat{x}, f(\hat{x}))$. Let x_0 be an initial guess for the root. To improve on the guess we construct the tangent line to the curve $y = f(x)$ that passes through the point $(x_0, f(x_0))$ on the curve. This tangent line satisfies the equation

$$y - f(x_0) = f'(x_0)(x - x_0).$$

The tangent line crosses the x-axis at the point

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)},$$

and we take x_1 as our improved estimate of the root $\hat{x}$. Now we repeat this procedure with x_1 to get an improved estimate x_2, and so on. Thus we have a sequence $\{x_n\}$ such that

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \qquad n = 0, 1, \cdots.$$

We need to give conditions that will guarantee that the sequence will converge to a root of $f(x)$, and will provide information about the rate of convergence.

To analyze this procedure we define an updating function $T(x)$ by

$$T(x) = x - \frac{f(x)}{f'(x)}.$$

We will not yet fix the domain D of this function, but it is clear that we must require $f'(x) \neq 0$ for all $x \in D$. Then $\hat{x}$ will be a fixed point of T, $(T(\hat{x}) = \hat{x})$ if and only if $f(\hat{x}) = 0$. To get the growth rate for the iteration we compute the derivative of $T(x)$:

$$T'(x) = \frac{f(x)f''(x)}{[f'(x)]^2}.$$

Since $T'(\hat{x}) = 0$, in the neighborhood of the root we will be able to select a decay constant $c < 1$, so the root is an attractive fixed point of T. In particular let $D = [\hat{x} - r, \hat{x} + r]$ where $|T'(x)| \leq c < 1$ for all $x \in D$. (If $f'(\hat{x}) \neq 0$ we can always find an r such that the inequality holds for the given constant c.) Then if $u, v \in D$, the Mean Value Theorem gives $T(u) - T(v) = T'(\tilde{u})(u - v)$ for some $\tilde{u} \in D$ between u and v. Thus $|T(u) - T(v)| \leq c|u - v|$ for all $u, v \in D$. In particular

$$|x_n - \hat{x}| \leq c|x_{n-1} - \hat{x}| \leq \cdots \leq c^n|x_0 - \hat{x}|.$$

Thus if $x_0 \in D$ then so are all of the $x_n \in D$ and $x_n \to \hat{x}$ as $n \to \infty$.

The convergence of the Newton algorithm is actually much faster than indicated from this analysis. This is due to the fact that $T'(\hat{x}) = 0$. We can, indeed, prove quadratic convergence. Suppose we can find a finite positive numbers A, B such that $B > |f''(x)|$ for all $x \in D$, and $A < |f'(x)|$ for all $x \in D$, and set $C = B/A$. By the Mean Value Theorem there is a point $\tilde{x}_n \in D$, between $\hat{x}$ and x_n, such that

$$f(x_n) = f(x_n) - f(\hat{x}) = f'(\tilde{x}_n)(x_n - \hat{x}),$$

so $x_n - \hat{x} = f(x_n)/f'(\tilde{x}_n)$. Furthermore, the Mean Value Theorem applied to $f'(x)$ yields a point $\breve{x}_n$ between x_n and $\tilde{x}_n$ such that

$$f'(x_n) - f'(\tilde{x}_n) = f''(\breve{x}_n)(x_n - \tilde{x}_n).$$

Then

$$|x_{n+1} - \hat{x}| = |(x_{n+1} - x_n) + (x_n - \hat{x})| = |\frac{f(x_n)}{f'(\tilde{x}_n)} - \frac{f(x_n)}{f'(x_n)}|$$

$$= |\frac{f(x_n)}{f'(x_n)f'(\tilde{x}_n)}(f'(x_n) - f'(\tilde{x}_n))| = |\frac{x_n - \hat{x}}{f'(x_n)}(f'(x_n) - f'(\tilde{x}_n))|$$

$$= |(x_n - \tilde{x}_n)(x_n - \hat{x})\frac{f''(\breve{x}_n)}{f'(x_n)}| \leq C|x_n - \hat{x}|^2.$$

Thus $|x_{n+1} - \hat{x}| \leq C|x_n - \hat{x}|^2$ and the convergence is quadratic. This means that the number of digits of accuracy in our approximation roughly doubles with each iteration.

Example 1 *We approximate $\sqrt{7}$ by using Newton's method to find the positive root of the function*

$$f(x) = x^2 - 7.$$

Here the iteration step is given by the function

$$T(x) = x - \frac{f(x)}{f'(x)} = \frac{x}{2} + \frac{7}{2x}.$$

Let $D = \{x : 2 \leq x \leq 7\}$ and choose the intial approximation $x_0 = 2$. Note that

$$\left| \frac{f(x) f''(x)}{(f'(x))^2} \right| = \left| \frac{1}{2} - \frac{7}{2x^2} \right| \leq \frac{3}{8} = c$$

for $x \in D$. Thus

$$|x_n - \sqrt{2}| \leq \frac{3}{8} |x_{n-1} - \sqrt{7}| \leq \cdots \leq (\frac{3}{8})^n |2 - \sqrt{7}| \to 0$$

as $n \to \infty$. The rate of covergence is much faster than this, however. Indeed $|f''(x)| = 2 = B$ and $|f'(x)| = |2x| \geq 4 = A$ for all $x \in D$. Thus, setting $C = \frac{B}{A} = \frac{1}{2}$, we have

$$|x_{n+1} - \sqrt{7}| \leq \frac{1}{2} |x_n - \sqrt{7}|^2, \quad n = 0, 1, \cdots,$$

and the number of guaranteed digits of accuracy more than doubles with each iteration. Indeed, we have (computing the first 10 digits)

$$\begin{aligned} x_0 &= 2 \\ x_1 &= 2.75 \\ x_2 &= 2.647727273 \\ x_3 &= 2.645752048 \\ x_4 &= 2.645751311 \end{aligned}$$

Here, x_4 is correct to more than 10 digits (if we calculated to that many decimal places) and $(x_4)^2 = 7.000000000$. Since x_3 has 6 digits accuracy, x_5 would have about 24 digits accuracy.

Example 2 *We approximate $\sqrt{7}$ by using the midpoint method to find the positive root of the function*

$$f(x) = x^2 - 7.$$

The method is symplicity itself. As in the last example, we start with the initial guess $x_0 = 2$ *and find* $f(2) = -3 < 0$. *For our next guess we take* $x_1 = 3$ *and find* $f(3) = 2 > 0$. *Note that* f *is a continuous function, so by the Intermediate Value Theorem it must have a root in the interval* $(2, 3)$, *of length 1. Now choose the midpoint of this interval:* $x_2 = (x_0 + x_1)/2 = 2.5$. *We see that* $f(x_2) = -.75$, *so by the Intermediate Value Theorem there must be a root of* f *in the interval* $(2.5, 3)$ *of length* $1/2$. *The midpoint of this interval is* $x_3 = 2.75$. *Now* $f(2.75) = .5625$, *so a root must lie in the interval* $(2.5, 2.75)$ *of length* $1/2^2$ *The midpoint of this interval is* $x_4 = 2.625$. *Since* $f(2.625) = -.109375$, *the root must lie in the interval* $(2.625, 2.75)$ *of length* $1/2^3$. *Continuing in this way, for each* $n > 0$ *we get an approximation* x_n *of* $\sqrt{7}$ *with accuracy* $1/2^{n-1}$. *This method is much simpler to implement than the Newton method, since we don't have to compute a derivative at each step, and it works for continuous functions that may not be differentiable at some points. However the rate of convergence is only linear, i.e., the error is cut in half at each step, whereas the Newton method has a quadratic convergence rate. Indeed for the midpoint method our approximations are*

$$\begin{aligned} x_0 &= 2 \\ x_1 &= 3 \\ x_2 &= 2.5 \\ x_3 &= 2.75 \\ x_4 &= 2.625 \\ x_5 &= 2.6875 \end{aligned}$$

and the convergence is very slow.