Quadratic Convergence of Newton's Method

Michael Overton, Numerical Computing, Spring 2017

The quadratic convergence rate of Newton's Method is not given in A&G, except as Exercise 3.9. However, it's not so obvious how to derive it, even though the proof of quadratic convergence (assuming convergence takes place) is fairly simple and may be found in many books. Here it is. Let f be a real-valued function of one real variable.

Theorem. Assume that f is twice continuously differentiable on an open interval (a, b) and that there exists $x^* \in (a, b)$ with $f'(x^*) \neq 0$. Define Newton's method by the sequence

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \qquad k = 1, 2, \ldots$$

Assume also that x_k converges to x^* as $k \to \infty$. Then, for k sufficiently large,

$$|x_{k+1} - x^*| \leq M|x_k - x^*|^2 \qquad \text{if } M > \frac{|f''(x^*)|}{2|f'(x^*)|}.$$

Thus, x_k converges to x^* *quadratically* (A&G, p. 52).

Proof. Let $e_k = x_k - x^*$, so that $x_k - e_k = x^*$. By Taylor's Theorem (A&G, Chap. 1, p. 5), setting $x = x_k$ and $h = -e_k$, we have

$$f(x_k - e_k) = f(x_k) - e_k f'(x_k) + \frac{(e_k)^2}{2} f''(\xi_k)$$

for some ξ_k between x_k and x^*. Since $x_k - e_k = x^*$ and $f(x^*) = 0$, we have

$$0 = f(x_k) - (x_k - x^*)f'(x_k) + \frac{(e_k)^2}{2} f''(\xi_k).$$

Since the derivative of f is continuous with $f'(x^*) \neq 0$, we have $f'(x_k) \neq 0$ as long as x_k is close enough to x^*. So we can divide by $f'(x_k)$ to give

$$0 = \frac{f(x_k)}{f'(x_k)} - (x_k - x^*) + \frac{(e_k)^2 f''(\xi_k)}{2f'(x_k)},$$

which, by the definition of Newton's method, gives

$$x_{k+1} - x^* = \frac{(e_k)^2 f''(\xi_k)}{2f'(x_k)}.$$

So

$$|x_{k+1} - x^*| \leq \frac{|f''(\xi_k)|}{2|f'(x_k)|} |x_k - x^*|^2.$$

By continuity, $f'(x_k)$ converges to $f'(x^*)$ and, since ξ_k is between x_k and x^*, ξ_k converges to x^* and hence $f''(\xi_k)$ converges to $f''(x^*)$, so, for large enough k,

$$|x_{k+1} - x^*| \leq M|x_k - x^*|^2 \qquad \text{if } M > \frac{|f''(x^*)|}{2|f'(x^*)|}.$$

In fact, it can be shown without assuming that x_k converges to x^*, that there exists $\delta > 0$ such that, if $|x_0 - x^*| \leq \delta$, then x_k converges to x^*, and hence from the above argument that the convergence rate is quadratic, but this requires a more complicated argument by induction.