2.7 Newton's Method for Nonlinear Equations

Let us now consider methods for solving

$$f(x) = 0.$$

We first consider the one-dimensional case where x is a scalar and f is a real-valued function. Later we will look at the n-dimensional case where $x = (x_1, \ldots, x_n)^T$ and $f(x) = (f_1(x), \ldots, f_n(x))^T$. Note that both x and $f(x)$ are vectors of the same length n. Throughout this section we assume that the function f has two continuous derivatives.

If $f(x)$ is a linear function, it is possible to find a solution if the system is nonsingular. The cost of finding the solution is predictable—it is the cost of applying Gaussian elimination. Except for a few isolated special cases, such as quadratic equations in one variable, in the nonlinear case it is not possible to guarantee that a solution can be found, nor is it possible to predict the cost of finding a solution. However, the situation is not totally bleak. There are effective algorithms that work much of the time, and that are efficient on a wide variety of problems. They are based on solving a *sequence* of linear equations. As a result, if the function f is linear, they can be as efficient as the techniques for linear systems. Also, we can apply our knowledge about linear systems in the nonlinear case.

The methods that we will discuss are based on Newton's method. Given an estimate of the solution x_k, the function f is approximated by the linear function consisting of the first two terms of the Taylor series for the function f at the point x_k. The resulting linear system is then solved to obtain a new estimate of the solution x_{k+1}.

To derive the formulas for Newton's method, we first write out the Taylor series for the function f at the point x_k:

$$f(x_k + p) \approx f(x_k) + pf'(x_k).$$

If $f'(x_k) \neq 0$, then we can solve the equation

$$f(x_*) \approx f(x_k) + pf'(x_k) = 0$$

for p to obtain

$$p = -f(x_k)/f'(x_k).$$

The new estimate of the solution is then $x_{k+1} = x_k + p$ or

$$x_{k+1} = x_k - f(x_k)/f'(x_k).$$

This is the formula for Newton's method.

Example 2.5 (Newton's Method)**.** As an example, consider the one-dimensional problem

$$f(x) = 7x^4 + 3x^3 + 2x^2 + 9x + 4 = 0.$$

Then

$$f'(x) = 28x^3 + 9x^2 + 4x + 9$$

and the formula for Newton's method is

$$x_{k+1} = x_k - \frac{7x_k^4 + 3x_k^3 + 2x_k^2 + 9x_k + 4}{28x_k^3 + 9x_k^2 + 4x_k + 9}.$$

Table 2.1. *Newton's method for a one-dimensional problem.*

k	x_k	$f(x_k)$	$\|x_k - x_*\|$
0	0	4×10^{0}	5×10^{-1}
1	−0.4444444444444444	4×10^{-1}	7×10^{-2}
2	−0.5063255748934088	3×10^{-2}	5×10^{-3}
3	−0.5110092428604380	2×10^{-4}	3×10^{-5}
4	−0.5110417864454134	9×10^{-9}	2×10^{-9}
5	−0.5110417880368663	0	0

If we start with the initial guess $x_0 = 0$, then

$$\begin{aligned} x_1 &= x_0 - \frac{7x_0^4 + 3x_0^3 + 2x_0^2 + 9x_0 + 4}{28x_0^3 + 9x_0^2 + 4x_0 + 9} \\ &= 0 - \frac{7 \times 0^4 + 3 \times 0^3 + 2 \times 0^2 + 9 \times 0 + 4}{28 \times 0^3 + 9 \times 0^2 + 4 \times 0 + 9} \\ &= 0 - \frac{4}{9} = -4/9 = -0.4444\ldots. \end{aligned}$$

At the next iteration we substitute $x_1 = -4/9$ into the formula for Newton's method and obtain $x_2 \approx -0.5063$. The complete iteration is given in Table 2.1. ∎

Newton's method corresponds to approximating the function f by its tangent line at the point x_k. The point where the tangent line crosses the x-axis (i.e., a zero of the tangent line) is taken as the new estimate of the solution. This geometric interpretation is illustrated in Figure 2.10.

The performance of Newton's method in Example 2.5 is considered to be typical for this method. It converges rapidly and, once x_k is close to the solution x_*, the error is approximately squared at every iteration. It has a quadratic rate of convergence as we now show.

It is not difficult to analyze the convergence of Newton's method using the Taylor series. Define the error in x_k by $e_k = x_k - x_*$. Using the remainder form of the Taylor series:

$$0 = f(x_*) = f(x_k - e_k) = f(x_k) - e_k f'(x_k) + \tfrac{1}{2} e_k^2 f''(\xi).$$

Dividing by $f'(x_k)$ and rearranging gives

$$e_k - \frac{f(x_k)}{f'(x_k)} = \frac{1}{2} e_k^2 \frac{f''(\xi)}{f'(x_k)}.$$

Since $e_k = x_k - x_*$ we obtain

$$x_k - \frac{f(x_k)}{f'(x_k)} - x_* = \frac{1}{2}(x_k - x_*)^2 \frac{f''(\xi)}{f'(x_k)},$$

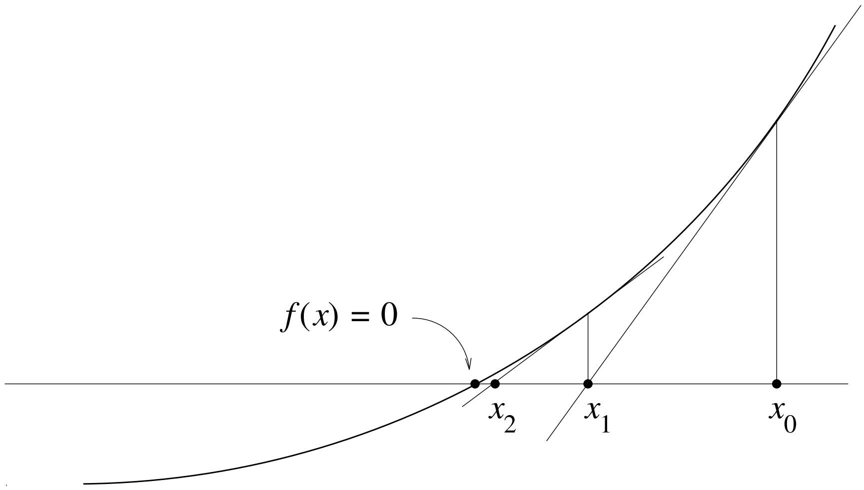

Figure 2.10. *Newton's method—geometric interpretation.*

which is the same as

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

If the sequence $\{ x_k \}$ converges, then $\xi \to x_*$, and hence when x_k is sufficiently close to x_*,

$$x_{k+1} - x_* \approx \frac{1}{2} \left(\frac{f''(x_*)}{f'(x_*)} \right) (x_k - x_*)^2$$

indicating that the error in x_k is approximately squared at every iteration, assuming that the rate constant $\frac{1}{2} f''(x_*)/f'(x_*)$ is not ridiculously large or small. These results are summarized in the following theorem.

Theorem 2.6 (Convergence of Newton's Method)**.** *Assume that the function $f(x)$ has two continuous derivatives. Let x_* be a zero of f with $f'(x_*) \neq 0$. If $|x_0 - x_*|$ is sufficiently small, then the sequence defined by*

$$x_{k+1} = x_k - f(x_k)/f'(x_k)$$

converges quadratically to x_* with rate constant

$$C = |f''(x_*)/2f'(x_*)|.$$

Proof. See the Exercises. □

Example 2.5 also shows that the function values $f(x_k)$ converge quadratically to zero. This also follows from the Taylor series:

$$0 = f(x_*) = f(x_k + e_k) = f(x_k) + e_k f'(\xi).$$

This can be rearranged to obtain

$$f(x_k) = -e_k f'(\xi) = -f'(\xi)(x_* - x_k)$$

so that $f(x_k)$ is proportional to $(x_* - x_k)$. Hence they converge at the same rate if $f'(x_*) \neq 0$.

In the argument above we have assumed that $\{ f'(x_k) \}$ and $f'(x_*)$ are all nonzero. If $f'(x_k) = 0$ for some k, then Newton's method fails (there is a division by zero in the formula). Geometrically this means that the tangent line is horizontal, parallel to the x-axis, and so it does not have a zero. If on the other hand $f'(x_k) \neq 0$ for all k, $f''(x_*) \neq 0$, but $f'(x_*) = 0$, then the coefficient in the convergence analysis

$$\frac{f''(\xi)}{2f'(x_k)}$$

tends to infinity, and the algorithm does not have a quadratic rate of convergence. If f is a polynomial, this corresponds to f having a multiple zero at the point x_*; this case is illustrated in Example 2.7.

Example 2.7 (Newton's Method; $f'(x_*) = 0$)**.** We now apply Newton's method to the example

$$\begin{aligned} f(x) &= x^4 - 7x^3 + 17x^2 - 17x + 6 \\ &= (x-1)^2(x-2)(x-3) = 0. \end{aligned}$$

This function has a multiple zero at $x_* = 1$ and at this point $f(x_*) = f'(x_*) = 0$. The derivative of f is

$$f'(x) = 4x^3 - 21x^2 + 34x - 17$$

and the formula for Newton's method is

$$x_{k+1} = x_k - \frac{x^4 - 7x^3 + 17x^2 - 17x + 6}{4x^3 - 21x^2 + 34x - 17}.$$

If we start with the initial guess $x_0 = 1.1$, then the method converges to $x_* = 1$ at a linear rate, whereas if we start with $x_0 = 2.1$, then the method converges to $x_* = 2$ at a quadratic rate. The results for these iterations are given in Tables 2.2 and 2.3. (In the final lines of both tables the function value $f(x_k)$ is zero; this is the value calculated by the computer and is a side effect of using finite-precision arithmetic.) ∎

In Example 2.7 the slow convergence only occurs when the method converges to a solution where $f'(x_*) = 0$. Quadratic convergence is obtained at the other roots, where $f'(x_*) \neq 0$.

It should also be noticed that the accuracy of the solution was worse at a multiple root. This too can be explained by the Taylor series, although this time we expand about the point x_*:

$$f(x_k) = f(x_* + e_k) = f(x_*) + e_k f'(x_*) + \tfrac{1}{2} e_k^2 f''(\xi).$$

At the solution, $f(x_*) = 0$, and since this is assumed to be a multiple zero, $f'(x_*) = 0$ as well. Hence

$$f(x_k) = \tfrac{1}{2} e_k^2 f''(\xi) = (\tfrac{1}{2} f''(\xi))(x_k - x_*)^2.$$