Let us revisit Newton's method once again: The equality we showed previously

$$g'(x) = \frac{f(x)f''(x)}{[f'(x)]^2}$$

can give us some insights about certain cases, where convergence is more likely, and others where convergence may be at risk:

- If f'' is *small*, $g'(x)$ will also tend to be small. In the limit case where $f''(x) = 0$, convergence is instantaneous; of course, this is of limited interest because it would imply that the equation of interest is in fact linear, or $f(x) = ax + b$. However, when $f''(x) \approx 0$, we can expect very rapid convergence.
- If $f'(x)$ is large, convergence will typically occur more easily. Of course, sometimes this fact coincides with f'' being large, in which case the two factors compete or cancel one another.
- Another consequence is that, when $f'(x) \approx 0$ (i.e., the plot of f is mostly "flat"), convergence will be less certain. Compare this with our intuitive graphical explanation of "flat" tangents in Newton's method.

Order of convergence

We previously used the hypothesis that g (in the fixed point iteration $x_{k+1} = g(x_k)$ is a contraction to show that $|x_k - a| \stackrel{k\to\infty}{\longrightarrow} 0$. Remember that the quantity

$$e = x_{\text{approximate}} - x_{\text{exact}}$$

was previously defined as the (absolute) error. In this case, let us define $e_k = x_k - a$ (a is the solution $f(a) = 0$) as the error after the k-th iteration. If g is a contraction, we have

$$|e_{k+1}| = |x_{k+1} - a| = |g(x_k) - g(a)| \leq L|x_k - a| = L|e_k|$$

Since $L < 1$ the error shrinks at least by a constant factor at each iteration.

In some cases we can do even better. Remember the following theorem:

Theorem (Taylor's formula) : If a function g is k-times differentiable, then:

$$g(y) = g(x) + g'(x)(y-x) + \frac{g''(x)}{2!}(y-x)^2 + \frac{g'''(x)}{3!}(y-x)^3 + \cdots$$

$$\cdots + \frac{g^{(k-1)}(x)}{(k-1)!}(y-x)^{k-1} + \frac{g^{(k)}(c)}{k!}(y-x)^k$$

for some number c between x and y.

For $k = 1$ we simply obtain the mean value theorem

$$g(y) = g(x) + g'(c)(y - x) \Leftrightarrow \frac{g(y) - g(x)}{y - x} = g'(c)$$

(for some c between x and y), which we used before to show that $|g'(x)| \leq L < 1$ implies that g is a contraction.

We will not use the theorem in the case $k = 2$:

$$g(y) = g(x) + g'(x)(y - x) + \frac{g''(c)}{2}(y - x)^2 \quad \text{for some } c \text{ between } x \& y \tag{3}$$

Let $g = x - f(x)/f'(x)$ (as in Newton's method). Now, let us make the following substitutions in the equation above:

$$x \leftarrow a \text{ (the solution), and } \; y \leftarrow x_k$$

If $f'(a) \neq 0$ and $f''(a)$ is defined, then

$$g'(a) = \frac{\overbrace{f(a)}^{=0} f''(a)}{[f'(a)]^2} = 0$$

Thus equation (3) becomes

$$g(x_k) = g(a) + \frac{g''(c)}{2}(x_k - a)^2$$

$$\Rightarrow x_{k+1} = a + \frac{g''(c)}{2}(x_k - a)^2$$

$$\Rightarrow |x_{k+1} - a| = \left|\frac{g''(c)}{2}\right| |x_k - a|^2$$

$$\Rightarrow |e_{k+1}| = C|e_k|^2 \quad \text{(note the exponent!)} \tag{4}$$

Where

$$C := \max \left|\frac{g''(x)}{2}\right|_{x \text{ between } a \text{ and } x_k}$$

Compare equation (4) with the general guarantee

$$|e_{k+1}| \leq L|e_k| \tag{5}$$

for contractions:

- Equation (5) depends on $L < 1$ to reduce the error. In equation (4), even if C is larger than one, if e_k is small enough then e_{k+1} will be reduced. Consider for example the case $C = 10$, $|e_k| = 10^{-3}$ which will guarantee $|e_{k+1}| \leq 10^{-5}$, and $|e_{k+2}| \leq 10^{-9}$ in the next iteration.

- Equation (5) implies that every iteration adds a fixed number (or, a fixed fraction) of correct significant digits. For example, if $L = 0.3$:

  $$|e_{k+2}| \leq 0.3|e_{k+1}| \leq 0.09|e_k|$$

  thus, we gain 1 significant digit every 2 iterations.

More generally, if an iterative scheme for solving $f(x) = 0$ can guarantee that

$$|e_{k+1}| \leq L|e_k|^d$$

then the exponent d (which can be a fractional number, too) is called the *order of convergence*. Specifically:

- If $d = 1$ we shall also require that $L < 1$ in order to guarantee that the error e_k is being reduced. In this case we say that the method exhibits *linear convergence*
- If $d > 1$ we no longer need $L < 1$ as a strict condition for convergence (although we need to be "close enough" to the solution to guarantee progress). This case is described as *superlinear* convergence. The case $d = 2$ is referred to as *quadratic* convergence, when $d = 3$ we talk about *cubic* convergence, and so on.

We previously saw that Newton's method converges quadratically. More generally, for the fixed point iteration $x_{k+1} = g(x_k)$, if we can show that $g'(a) = 0$ (a is the solution), then Taylor's 2nd order formula yields the same result as in equation (4), and the fixed point iteration converges quadratically.

Multiple roots

So far we have ignored the case $f'(a) = 0$. This case is typically described as a *multiple root* because if $f(x)$ is a polynomial and $f(a) = f'(a) = f''(a) = \cdots = f^{(k-1)}(a) = 0$ this would imply that $(x - a)^k$ is a factor of $f(x)$ (in other words, a is a root with multiplicity of k).

Let us now assume that $f(a) = f'(a) = f''(a) = \cdots = f^{(k-1)}(a) = 0$, but $f^{(k)}(a) = 0$. At first, it may appear that Newton's method would be inapplicable in this case, because the denominator $f'(x)$ becomes near-zero close to the solution. Consider, however, the example:

$$f(x) = (x - 1)^3(x + 2)$$

Newton's method would give:

$$g(x) = x - \frac{f(x)}{f'(x)} = x - \frac{(x-1)^3(x+2)}{3(x-1)^2(x+2) + (x-1)^3} = x - \frac{(x-1)(x+2)}{4x+5}$$

whose denominator remains non-zero near the solution $x = 1$. In fact we can show that g remains a contraction near a in this case.

Optional reading

Proof : Remember that

$$g'(x) = \frac{f(x)f''(x)}{[f'(x)]^2}$$

Taylor's theorem applied on $f(x)$ yields:

$$f(x) = \underbrace{f(a)}_{=0} + \underbrace{f'(a)}_{=0}(x-a) + \cdots + \underbrace{\frac{f^{(k-1)}(a)}{(k-1)!}}_{=0}(x-a)^{k-1} + \underbrace{\frac{f^{(k)}(c_1)}{k!}}_{\neq 0}(x-a)^k$$

Applying Taylor's formula on the *derivative* $f'(x)$ gives

$$f'(x) = \underbrace{f'(a)}_{=0} + \underbrace{f''(a)}_{=0}(x-a) + \cdots + \underbrace{\frac{f^{(k-1)}(a)}{(k-2)!}}_{=0}(x-a)^{k-2} + \underbrace{\frac{f^{(k)}(c_2)}{(k-1)!}}_{\neq 0}(x-a)^{k-1}$$

And, one more, on the second derivative $f''(x)$

$$f''(x) = \underbrace{f''(a)}_{=0} + \underbrace{f'''(a)}_{=0}(x-a) + \cdots + \underbrace{\frac{f^{(k-1)}(a)}{(k-3)!}}_{=0}(x-a)^{k-3} + \underbrace{\frac{f^{(k)}(c_3)}{(k-2)!}}_{\neq 0}(x-a)^{k-2}$$

Where c_1, c_2, c_3 are some numbers between x and a. Combining the last 3 equations, we get

$$g'(x) = \frac{\frac{f^{(k)}(c_1)}{k!}(x-a)^k \frac{f^{(k)}(c_3)}{(k-2)!}(x-a)^{k-2}}{\left[\frac{f^{(k)}(c_2)}{(k-1)!}(x-a)^{k-1}\right]^2} \xrightarrow{x\to a} \frac{\left[f^{(k)}(a)\right]^2}{\left[f^{(k)}(a)\right]^2} \frac{[(k-1)!]^2}{k!(k-2)!} = \frac{k-1}{k}.$$

Since $g'(a) = (k-1)/k < 1$, there is an interval $(a-\delta, a+\delta)$ where $g'(x) \leq L < 1$, and thus g is a contraction.

However, in all of these cases, convergence is limited to be *only linear*, since $g'(a) \neq 0$.

Caution : If we know, or suspect, that $f(x)$ may have a multiple root, then Newton's method is only safe (albeit slow) when we can write an analytic formula for $f(x)/f'(x)$ and perform any cancellations "on paper" to avoid division by zero. Otherwise, for example in the case where both f and f' are given as black-box computer functions, any roundoff error in the value of the numerator or denominator could cause severe instabilities, by dividing two (inaccurate) near-zero quantities.