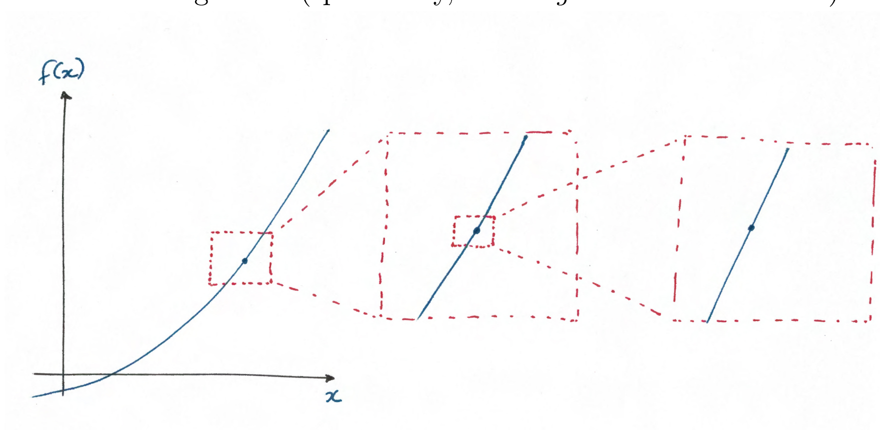
more application of this equation would yield $e_{k+2} < 10^{-12}$. Thus we see that, provided the iteration starts *close enough* to the solution, we not only converge to the desired value, but actually double the number of correct significant digits in each iteration. We defer the detailed proof until after we have introduced the more general method.

Lecture of: 29 Jan 2013

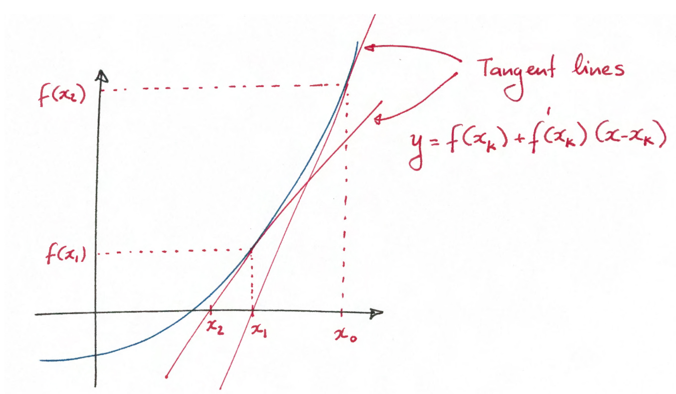
Newton's method

This example is a special case of an algorithm for solving nonlinear equations, known as Newton's method (also called the *Newton-Raphson* method). The general idea is as follows: If we "zoom" close enough to any smooth function, its graph looks more and more like a straight line (specifically, the *tangent* line to the curve).



Newton's method suggests: If after k iterations we have approximated the solution of $f(x) = 0$ (a general nonlinear equation) as x_k, then:

- Form the tangent line at $(x_k, f(x_k))$
- Select x_{k+1} as the intersection of the tangent line with the horizontal axis $(y = 0)$.



If $(x_n, y_n) = (x_n, f(x_n))$, the tangent line to the plot of $f(x)$ at (x_n, y_n) is:

$$y - y_n = \lambda(x - x_n), \quad \text{where } \lambda = f'(x_n) \text{ is the slope}$$

Thus the tangent line has equation $y - y_n = f'(x_n)(x - x_n)$. If we set $y = 0$ we get:

$$-f(x_n) = f'(x_n)(x - x_n) \Rightarrow x = x_n - \frac{f(x_n)}{f'(x_n)} := x_{n+1}$$

Ultimately, Newton's method reduces to: $\boxed{x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}}$

Our previous example (square root of a) is just an application of Newton's method to the nonlinear equation $f(x) = x^2 - a = 0$. Newton gives:

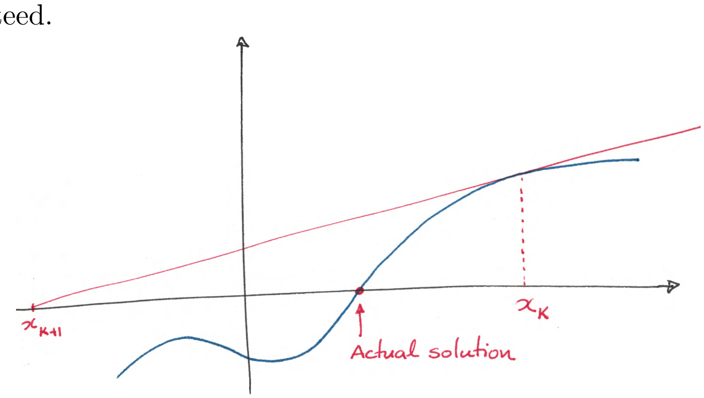
$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} = x_k - \frac{x_k^2 - a}{2x_k} = \frac{2x_k^2 - x_k^2 + a}{2x_k} = \frac{x_k^2 + a}{2x_k}$$

which is the same iteration we considered previously.

A few comments about Newton's method:

- It requires the function $f(x)$ to be not only continuous, but differentiable as well. We will later see variants that do not *explicitly* required knowledge of f'. This would be an important consideration if the formula for $f'(x)$ is significantly more complex, and expensive to evaluate than $f(x)$, or in the case we simply do not possess an analytic expression for f'; this could be the case if $f(x)$ is not given to us via an explicit formula, but only defined via a black-box computer function that computes its value.

- If we ever have an approximation x_k with $f'(x_k) \approx 0$, we should expect problems, especially if we are not close to a solution (we would be nearly dividing by zero). In such cases, the tangent line is almost (or exactly) horizontal, thus the next iterate can be a very remote value, and convergence may be far from guaranteed.



Fixed point iteration

Corresponding textbook chapter(s): §2.1

Newton's method is in itself a special case of a broader category of methods for solving nonlinear equations called *fixed point iteration* methods. Generally, if $f(x) = 0$ is the nonlinear equation we seek to solve, a fixed point iteration method proceeds as follows:

- Start with $x_0 = <$ *initial guess* $>$
- Iterate the sequence
$$x_{k+1} = g(x_k)$$
where $g(x)$ is a properly designed function for this purpose.

Following this method, we construct the sequence $x_0, x_1, x_2, \ldots, x_k, \ldots$ hoping that it will converge to a solution of $f(x) = 0$. The following questions arise at this point:

1. If this sequence converges, does it converge *to a solution* of $f(x) = 0$?
2. Is the iteration guaranteed to converge?
3. How fast does the iteration converge?
4. (Of practical concern) When do we stop iterating, and declare that we have obtained an acceptable approximation?

Note: The function $g(x)$ is related, but otherwise different than $f(x)$. In fact, there could be *many* different functions $g(x)$ that could be successfully used in a fixed point iteration of this form, all for solving the same equation $f(x) = 0$.

We start by addressing the first question: If the sequence $\{x_k\}$ does converge, can we ensure that it will converge to a solution of $f(x) = 0$?

Taking limits on $x_{k+1} = g(x_k)$, and assuming that (a) $\lim x_k = a$ and (b) the function g is continuous, we get:

$$\lim_{k\to\infty} x_{k+1} = \lim_{k\to\infty} g(x_k) \Rightarrow a = g(a)$$

That is, the limit of the fixed point iteration (if it exists) will satisfy the equation $a = g(a)$. The question is: can we guarantee that this limit value will actually be a solution to our original equation, i.e. $f(a) = 0$? The simplest way to ensure this property is to construct $g(x)$ such that

$$x = g(x) \ \text{ is mathematically equivalent to } \ f(x) = 0.$$

There are many ways to make this happen, e.g.

$$f(x) = 0 \Leftrightarrow x + f(x) = x \Leftrightarrow x = g(x), \quad \text{where} \quad g(x) := x + f(x)$$

or

$$f(x) = 0 \Leftrightarrow e^{-x} f(x) = 0 \Leftrightarrow e^{-x} f(x) + x^2 = x^2 \Leftrightarrow \frac{e^{-x} f(x) + x^2}{x} = x \Leftrightarrow$$

$$\Leftrightarrow g(x) = x, \quad \text{where} \quad g(x) := \frac{e^{-x} f(x) + x^2}{x}$$

or

$$f(x) = 0 \Leftrightarrow -\frac{f(x)}{f'(x)} = 0 \Leftrightarrow x - \frac{f(x)}{f'(x)} = x \Leftrightarrow g(x) = x, \quad \text{where} \quad g(x) := x - \frac{f(x)}{f'(x)}$$

The last example is exactly Newton's method; substituting the definition of $g(x)$ above into the iteration $x_{k+1} = g(x_k)$ yields the familiar Newton update equation. Thus we know that if Newton converges, it will be to a solution of $f(x) = 0$.

As we see, there are many possible ways to define the function $g(x)$, all of which guarantee at least one desirable property : *if* the iteration ever converges, the limit will be a solution of $f(x) = 0$. Unfortunately, not all of these different possibilities for $g(x)$ will lead to a method that actually converges! For example, consider the nonlinear equation $f(x) = x^2 - a = 0$ (Solution: $\pm\sqrt{a}$) and the function $g(x) = a/x$. We can easily verify that

$$x = g(x) = \frac{a}{x} \Leftrightarrow x^2 = a \Leftrightarrow x^2 - a = 0 = f(x)$$

However, the iteration $x_{k+1} = g(x_k) = a/x_k$ yields

$$x_1 = \frac{a}{x_0}, \quad x_2 = \frac{a}{x_1} = \frac{a}{a/x_0} = x_0$$

Thus, the sequence alternates forever between the values $x_0, x_1, x_0, x_1, \ldots$ regardless of the initial value. Other choices of $g(x)$ may also create divergent sequences, often regardless of the value of the initial guess.

Fortunately, there are ways to ensure that the sequence $\{x_k\}$ converges, by making an appropriate choice of $g(x)$. We will use the following definition:

Definition: A function $g(x)$ is called a *contraction in the interval* $[a, b]$, if there exists a number $L \in [0, 1)$ such that

$$|g(x) - g(y)| \leq L|x - y|$$

for any $x, y \in [a, b]$.

Examples:

- $g(x) = x/2$:
$$|g(x) - g(y)| = \frac{1}{2}|x - y|$$
for any $x, y \in \mathbf{R}$.

- $g(x) = x^2$, in $[0.1, 0.2]$:
$$|g(x) - g(y)| = |x^2 - y^2| = |x + y||x - y| \leq 0.4|x - y|$$
for $x, y \in [0.1, 0.2]$ (in this case this condition is essential!)

If we can establish that the function g in the fixed point iteration $x_{k+1} = g(x_k)$ is a contraction, we can show the following:

Theorem: Let a be the actual solution of $f(x) = 0$, and assume $|x_0 - a| < \delta$ (δ is an arbitrary positive number). If g is a contraction on $(a - \delta, a + \delta)$, the fixed point iteration is guaranteed to converge to a.

Any proofs printed in a shaded box need *not* be memorized for exams

Proof : Since a is the solution we have $a = g(a)$. Thus:

$$\begin{aligned}
|x_1 - a| &= |g(x_0) - g(a)| \leq L|x_0 - a| < L\delta \\
|x_2 - a| &= |g(x_1) - g(a)| \leq L|x_1 - a| < L^2\delta \\
&\vdots \\
|x_k - a| &< L^k\delta
\end{aligned}$$

Since $L < 1$ we have $\lim |x_k - a| = 0$, i.e. $x_k \to a$.

In some cases, it can be cumbersome to apply the definition directly to show that a given function g is a contraction. However, if we can compute the derivative $g'(x)$ we have a simpler criterion:

Theorem: If g is differentiable and a number $L \in [0, 1)$ exists such that $|g'(x)| \leq L$ for all $x \in [a, b]$, then g is a contraction on $[a, b]$.

Proof : Let $x, y \in [a, b]$ and, without loss of generality, assume $x < y$. The mean value theorem states that

$$\frac{g(x) - g(y)}{x - y} = g'(c) \quad \text{for some } c \in (x, y).$$

Now, if $|g'(x)| \le L$ for all $x \in [a, b]$, then regardless of the exact value of c we have

$$|g'(c)| \le L \Rightarrow \left| \frac{g(x) - g(y)}{x - y} \right| \le L \Rightarrow |g(x) - g(y)| \le L|x - y|$$

Examples:

- Let $g(x) = \sin(\frac{2x}{3})$. Then:

  $$|g'(x)| = \frac{2}{3} \left| \cos\left(\frac{2x}{3}\right) \right| \le \frac{2}{3} < 1$$

  thus g is a contraction.

Lecture of: 31 Jan 2013

- Let us try to apply the derivative criterion, to see if the function

  $$g(x) = x = f(x)/f'(x)$$

  which defines Newton's method is a contraction:

  $$g'(x) = 1 - \frac{f'(x)f'(x) - f(x)f''(x)}{[f'(x)]^2} = \frac{f(x)f''(x)}{[f'(x)]^2}$$

  Now let us assume that

  - $f(a) = 0$, i.e. a is a solution of $f(x) = 0$,
  - $f'(a) \neq 0$, and
  - f'' is *bounded* near a (for example, if f'' is continuous).

  Then

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

  This means that there is an interval $(a - \delta, a + \delta)$ where $|g'(x)|$ is *small* (since $\lim g'(x) = 0$). Specifically, we can find an $L < 1$ such that $|g'(x)| \le L$ when $|x - a| < \delta$. This means that g is a contraction on $(a - \delta, a + \delta)$, and if the initial guess also falls in that interval, the iteration is guaranteed to converge to the solution a.