

Notes: Rate of Convergence

Suppose we have a sequence of real numbers that converge to some point x :

$$x_0, x_1, x_2, \dots \rightarrow x$$

Is there a way to talk about how fast the numbers are converging to x ?

Definition: If

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - x|}{|x_n - x|} = \lambda < \infty$$

then the sequence converges *linearly* to x . The constant $\lambda > 0$ is called the asymptotic error.

Definition: If

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - x|}{|x_n - x|^2} = \lambda < \infty$$

(where $\lambda > 0$) then the sequence converges *quadratically* to x .

Definition: If

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - x|}{|x_n - x|^\alpha} = \lambda < \infty$$

then the sequence converges to x of order α . This is the definition we use to actually compute the rate of convergence- Typically, this means that we need to see if $\alpha = 1$ or 2 .

Examples:

1. Let $x_n = \frac{1}{n^k}$ for some fixed $k > 0$. Then we know this sequence converges to 0. Compute the rate of convergence:

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1}|}{|x_n|^\alpha} = \lim_{n \rightarrow \infty} \frac{\frac{1}{(n+1)^k}}{\frac{1}{n^{\alpha k}}} = \lim_{n \rightarrow \infty} \left(\frac{n^\alpha}{n+1} \right)^k$$

We get convergence for $\alpha = 1$, but not $\alpha = 2$, so this sequence converges **linearly** to 0.

This is interesting- Sequences like the following:

$$x_n = \frac{1}{n}, \quad x_n = \frac{1}{n^2}, \quad x_n = \frac{1}{n^3}, \dots$$

all converge **linearly** to zero.

2. Let $x_n = 10^{-2^n}$ (which converges to 0). Compute the rate of convergence:

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1}|}{|x_n|^\alpha} = \lim_{n \rightarrow \infty} \frac{10^{-2^{n+1}}}{10^{-2^n \cdot \alpha}}$$

It's easy to see that if $\alpha = 2$, we get convergence.

In fact, using this as a template sequence, we can see how fast a quadratically convergent sequence moves. Compute the terms of the sequence:

$$10^{-2}, 10^{-4}, 10^{-8}, 10^{-16}, 10^{-32}, 10^{-64}$$

so that after just 6 steps, we're already past machine ϵ . Algorithms that are quadratically convergent are highly prized!

0.1 Fixed Point Iteration

Now let's analyze the fixed point algorithm, $x_{n+1} = f(x_n)$ with fixed point r . We will see below that the key to the speed of convergence will be $f'(r)$.

Theorem (Convergence of Fixed Point Iteration): Let f be continuous on $[a, b]$ and f' be continuous on (a, b) . Furthermore, assume there exists $k < 1$ so that $|f'(x)| \leq k$ for all x in (a, b) .

- If $f'(r) \neq 0$, the sequence converges *linearly* to the fixed point.
- If $f'(r) = 0$, the sequence converges at least quadratically to the fixed point (this is sometimes called *superconvergence* in the dynamical systems literature).

Proof: This is more than we did in class- We will also prove that the sequence $x_n \rightarrow r$, which we do first. The following computations show this (they come from a repeated application of the Mean Value Theorem):

$$|x_1 - r| = |f(x_0) - f(r)| \leq |f'(x^*)||x_0 - r| \leq k|x_0 - r|$$

where x^* is between x_0 and r .

Continuing, we see that:

$$|x_n - r| \leq |f'(x^*)||x_{n-1} - r| \leq k|x_{n-1} - r| \leq k^n|x_0 - r|$$

with $0 \leq k < 1$, so that $k^n \rightarrow 0$ as $n \rightarrow \infty$. Therefore, $|x_n - r| \rightarrow 0$ as $n \rightarrow \infty$. An important note: While x^* depended on x_{n-1} , the value of k does not depend on n - The assumption that $|f'(x)| \leq k < 1$ was important for this step.

Now we show that the sequence converges *linearly* if $f'(r) \neq 0$:

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - r|}{|x_n - r|} = \lim_{n \rightarrow \infty} \frac{|f(x_n) - f(r)|}{|x_n - r|} = \lim_{n \rightarrow \infty} |f'(c_n)|$$

where c_n is between x_n and r . Note that, because $x_n \rightarrow r$, then $c_n \rightarrow r$ as well. Because f' is continuous on our interval that contains all the c_n 's and r ,

$$\lim_{n \rightarrow \infty} |f'(c_n)| = \left| f' \left(\lim_{n \rightarrow \infty} c_n \right) \right| = |f'(r)|$$

To get an idea about why the second case should give quadratic convergence, consider the Taylor expansion of f about the fixed point r :

$$f(x) = f(r) + f'(r)(x - r) + \frac{1}{2}f''(c)(x - r)^2$$

where c is in the interval between x and r .

Now, if $f'(r) = 0$, and rearranging terms, we see that:

$$\frac{f(x) - f(r)}{(x - r)^2} = \frac{f''(x^*)}{2}$$

Now to prove the quadratic convergence,

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - r|}{|x_n - r|^2} = \lim_{n \rightarrow \infty} \frac{|f(x_n) - f(r)|}{|x_n - r|^2} = \lim_{n \rightarrow \infty} \frac{|f''(c_n)|}{2}$$

where c_n is in the interval between x_n and r . And the proof is finished just as before (by assuming that f'' is continuous on the interval).

Before continuing, there is a property of continuous functions that we want to consider (used in the next section):

Lemma: Let f be continuous on some open interval I which contains a point a where $f(a) \neq 0$. Then there exists an interval about a (contained in I), $[a - \delta, a + \delta]$ such that $f(x) \neq 0$ for all x in $[a - \delta, a + \delta]$.

Proof: The proof is by contradiction- That is, if we want to prove that statement A implies statement B, assume A and not B. In this case, we assume that f is continuous, and for every δ -sized interval about a , there is at least one point so that $f(x) = 0$. If that is true, then let δ be any number, and choose $a_1 \neq a$ in the interval about a so that $f(a_1) = 0$. Next, let δ be half the distance between a and a_1 . In this interval, we can find $a_2 \neq a$ so that $f(a_2) = 0$. Next, we repeat by letting δ be half the distance from a_2 to a , and find a_3 . Repeat this process.

Now we have constructed a sequence of points, a_n , which converge to a . Furthermore, $f(a_n) = 0$ for all n , which means that:

$$\lim_{n \rightarrow \infty} f(a_n) = 0$$

But, since $a_n \rightarrow a$ and f is continuous at a , we must have:

$$\lim_{n \rightarrow \infty} f(a_n) = f\left(\lim_{n \rightarrow \infty} a_n\right) = f(a) \neq 0$$

This is a contradiction. Therefore, we conclude the opposite- That there must be an interval about a so that $f(x) \neq 0$ in the entire interval.

0.2 Newton's Method

As we will recall, Newton's Method was an algorithm to solve $g(x) = 0$ (this is not yet a fixed point algorithm). Given an initial x_0 , we compute a sequence of points defined via:

$$x_{n+1} = x_n - \frac{g(x_n)}{g'(x_n)}$$

Some questions we'd like to answer:

- Under what conditions will Newton's Method converge?
- If the sequence does converge, at what rate does it converge?
- Let r be the root of g . What happens if $g'(r) = 0$?

As you might expect, the answer to the first question is not as exact as it was for Fixed Point Iteration:

Theorem (Convergence of Newton's Method): Let g be twice continuously differentiable on the interval (a, b) . If $r \in (a, b)$ such that $g(r) = 0$ and $g'(r) \neq 0$, then there exists $\delta > 0$ such that Newton's Method will converge if started in the interval $[r - \delta, r + \delta]$. In this case, the sequence converges quadratically.

We make one observation to begin: Newton's Method is a form of Fixed Point iteration:

$$x_{n+1} = F(x_n) \quad \text{where} \quad F(x) = x - \frac{g(x)}{g'(x)}$$

and the convergence of fixed point iteration depended on the derivative of F :

$$F'(x) = 1 - \frac{g'(x)g'(x) - g(x)g''(x)}{(g'(x))^2} = \frac{g(x)g''(x)}{(g'(x))^2}$$

we also see that $F'(r) = 0$.

To be able to use the Fixed Point Theorem (Item 3 above), we would need to show that F satisfies those conditions- Namely, that there is an interval (a, b) about r where F' is continuous and $|F'(x)| < 1$. Let's see how our assumptions create the appropriate δ so that we guarantee the convergence of Newton's Method. Here are some observations:

- (Continuity of F and F') F will be continuous if g and g' are continuous, and $g'(x) \neq 0$.

Now we have no problem with g - It is continuous on (a, b) .

Since $g'(r) \neq 0$, and g' is continuous, we can guarantee that there exists δ_1 so that $g'(x) \neq 0$ for all x in $[r - \delta_1, r + \delta_1]$ (note that this interval needs to be contained in (a, b) , so make the choice of δ_1 appropriately). In this interval, F is continuous. Since we also assumed that g'' is continuous on (a, b) , then F' will also be continuous on $[r - \delta_1, r + \delta_1]$

- (The size of $F'(x)$) We know that $F'(r) = 0$, and we know that $F'(x)$ is continuous on $[r - \delta_1, r + \delta_1]$. By the continuity of F' , there must exist $0 < \delta_2 \leq \delta_1$ so that $|F'(x)| < 1$ for all x in $[r - \delta_2, r + \delta_2]$. The value δ_2 is the δ in our Theorem.

By our two previous observations, we can apply the Fixed Point Convergence theorem to the function F - That proves both convergence and (at least) quadratic convergence.

0.2.1 EXERCISE:

Prove that, if $g(r) = 0$ and g is continuous on an interval about r , then there exists δ so that $|g(x)| \leq k < 1$ for all x in $(r - \delta, r + \delta)$. (HINT: Follow the proof of the Lemma)

0.2.2 EXERCISE:

Give an interval on which we can guarantee the convergence of Newton's Method, if $g(x) = x^2 - 1$ (give the interval about the root $r = 1$).

0.2.3 Failure to Converge Quadratically:

Consider $g(x) = x^2$. Will Newton's Method converge quadratically to the root $x = 0$? First, we convert to the Fixed Point Iteration method and see what we can conclude:

$$F(x) = x - \frac{g(x)}{g'(x)} = x - \frac{x^2}{2x} = x - \frac{x}{2} = \frac{1}{2}x$$

In this case, $F'(0) = \frac{1}{2} \neq 0$, so Newton's Method will converge, but only linearly.

This happened because there was a multiple root at $x = 0$:

Definition: A function $f(x)$ has a root of multiplicity m at $x = a$ if:

$$f(a) = 0, f'(a) = 0, f''(a) = 0, \dots, f^{(m)}(a) = 0, \quad \text{but } f^{(m+1)}(a) \neq 0$$

Be sure to keep the conclusions of the Fixed Point Method and Newton's Method distinct:

- In Fixed Point Iteration, if $F'(r) = 0$, we get at least quadratic convergence. If $F'(r) \neq 0$, we get linear convergence.
- In Newton's Method, if $g'(r) \neq 0$, we get quadratic convergence, and if $g'(r) = 0$, we get only linear convergence.

0.2.4 Exercise:

The function $f(x) = x^3 - 4x$ has a root at $r = 2$. If the error $e_i = x_i - r$ after 4 steps of Newton's Method is $e_4 = 10^{-6}$, estimate e_5 . (Hint: What is the limit of e_{n+1}/e_n as $n \rightarrow \infty$?)

0.2.5 Practical Notes for Newton's Method:

If you believe your function has a "simple" root (that is, $g(r) = 0$ and $g'(r) \neq 0$), and you have a reasonable starting guess, x_0 , then Newton's Method is the method of choice. In practice, we will not know if we have a simple root, and we may not have a very good starting estimate. In that case, you might use something like Bisection to get you started, then switch to Newton's Method once you're close. We'll see this in a method later on.

1 Methods without Derivatives

There are methods other than bisection if the function's derivative is not known (or not easily computed). We discuss the ideas briefly below:

- **Secant Method** Rather than using the x -intercept of the tangent line (as in Newton), we begin with two points, x_0 and x_1 , then use the x -intercept of the secant line as our new estimate, x_2 . In this method, we do not need to start with the root bracketed (as in bisection), and in fact the root may not be bracketed.
- **Method of False Position** Similar to bisection, but the midpoint is replaced by the x -intercept of the line between $(a, f(a))$ and $(b, f(b))$. The new interval is chosen so that the root is still bracketed. Note that this method may not do as well as bisection, which halves the interval at each step.

- **Muller’s Method** This is a generalization of the Secant Method in that it uses three points, x_0, x_1, x_2 and uses a parabola instead of a line to get the next point. Kind of an interesting idea, and we’ll discuss finding that parabola later.

These methods work fairly well under the right circumstances. The next method we will do in some detail, as it forms the basis of Matlab’s `fzero` function.

1.0.6 Exercise:

From the description of the Method of False Position, give an explicit algorithm for finding a root of f . You can assume that you are given $[a, b]$ such that $f(a)f(b) < 0$. Apply the algorithm to $f(x) = x^3 - 2x^2 + \frac{3}{2}x$ on $[-1, 1]$. Take special note of the sizes of the intervals that you create.

1.1 Inverse Quadratic Interpolation (IQI)

In the IQI algorithm, we construct the quadratic polynomial $x = p(y)$ rather than $y = p(x)$. This has the major benefit that (i) the x -intercept is very easy to find and (ii) we are always guaranteed that the parabola passes through the x -axis once.

As we know, three points defines a parabola. These three points:

$$(a, A), (b, B), (c, C)$$

can be used to construct the following quadratic polynomial:

$$P(y) = a \frac{(y - B)(y - C)}{(A - B)(A - C)} + b \frac{(y - A)(y - C)}{(B - A)(B - C)} + c \frac{(y - A)(y - B)}{(C - A)(C - B)}$$

(This is a Lagrange polynomial, and we’ll discuss them later) Note that $P(A) = a$, $P(B) = b$, and $P(C) = c$ (that is what it means to say that the polynomial *interpolates* the data points).

So, given $a = x_i$, $b = x_{i+1}$ and $c = x_{i+2}$, then x_{i+3} is computed as the x -intercept of $P(y)$ (computed as $P(0)$):

$$x_{i+3} = P(0) = \frac{aBC}{(A - B)(A - C)} + \frac{bAC}{(B - A)(B - C)} + \frac{cAB}{(C - A)(C - B)}$$

1.2 Brent’s Method

Brent’s method is the actual workhorse in root-finding without the use of derivatives. It is what Matlab uses in `fzero`¹.

Brent’s Method combines root bracketing, bisection and IQI to converge from the neighborhood of a zero crossing. To see what Matlab actually does, enter the following in the command window: `type fzero`.

Most of what you see is “error handling”- That is, in a professional piece of software, you have to try to anticipate all the problems that might come up. Here are a few problems associated with `fzero`:

¹Matlab also has a `roots` command specifically for polynomials, but this technique uses eigenvalues to solve for the roots- We’ll see that later.

- How is the function being defined? It can be entered as an inline function (like our examples), or as the name of an M-file.
- Did the user input an appropriate bracket for the root, or an initial guess, from which we have to construct an initial bracket? Is the initial guess appropriate (for example, did the user input -1 for \sqrt{x})?
- How much output does the user want? The default is for Matlab to stay quiet until the final result, but you can have it output details about what it is doing. You might try this example:

```
fzero('x.^3+x-1',1,optimset('Display','iter'))
```

You can see `fzero` trying to do an initial bracketing before the algorithm proceeds.

- How should the algorithm stop, and what information should be presented? Lots of options here.