Theorem 1.7.2. *Let, as usual, $g \in C[a,b]$ and suppose that $g(x) \in [a,b]$ for all $x \in [a,b]$. Further suppose that the derivative exist on (a,b) and*

$$g^{'}(x) \leq c < 1 \quad \textit{for all} \quad x \in (a,b)$$

If we take any initial value $x_0 \in (a,b)$ then the sequence arising from

$$x_n = g(x_{n-1})$$

converges to the unique fixed point $r \in [a,b]$.

Proof: We already know that there must exist a unique fixed point in $[a,b]$ from our previous theorem. Note now that from the assumption that $g(x) \in [a,b]$ for all $x \in [a,b]$ we must also have that all the x_n's are also in $[a,b]$. Using (1.3) and the mean value theorem we then must have,

$$|x_n - r| = |g(x_{n-1}) - g(r)| = |g^{'}(\xi)||x_{n-1} - r| \leq c|x_{n-1} - r|$$

for some $\xi \in (a,b)$. We now keep reapplying this inequality on the above and obtain,

$$|x_n - r| \leq c|x_{n-1} - r| \leq c^2|x_{n-2} - r| < \cdots \leq c^{n-1}|x_1 - r| \leq c^n|x_0 - r| \tag{1.5}$$

However, $c < 1$. Thus,

$$\lim_{n\to\infty} |x_n - r| \leq \lim_{n\to\infty} c^n|x_0 - r| = 0$$

That is the sequence of $\{x_n\}_{n=1}^{\infty}$ converges to the true solution r.

1.8 Error estimates for fixed point iteration

We can now use this theorem to obtain some very useful results about the error of our approximations regarding the fixed point algorithm. Both of the inequalities below are a-priori type estimates.

Corollary 1.8.1. *Suppose g as in Theorem (1.7.2) above. Then the error of our approximating the fixed point r of g can be estimated by either of the following two formulas:*

$$|x_n - r| \leq c^n \max\{x_0 - a, b - x_0\} \tag{1.6}$$

$$|x_n - r| \leq \frac{c^n}{1-c}|x_1 - x_0| \tag{1.7}$$

for all $n \geq 1$.

Proof: Since r is in $[a,b]$ the first formula (1.6) is clear from inequality (2.9).

To prove the second formula (1.7) we need some more work. Note that, as in Theorem (1.7.2), the following is true,

$$|x_{n+1} - x_n| = |g(x_n) - g(x_{n-1})| \leq c|x_n - x_{n-1}| < \cdots \leq c^n|x_1 - x_0|$$

Therefore for any $m > n \geq 1$ by adding and subtracting similar terms the following must be true,

$$|x_m - x_n| = |x_m - x_{m-1} + x_{m-1} - \cdots - x_{n+1} + x_{n+1} - x_n|$$

But the right hand side of this is clearly less than,

$$\leq |x_m - x_{m-1}| + |x_{m-1} - x_{m-2}| + \cdots + |x_{n+1} - x_n|$$

Or further,

$$\leq c^{m-1}|x_1 - x_0| + c^{m-2}|x_1 - x_0| + \cdots + c^n|x_1 - x_0|$$

and finally,

$$= c^n(1 + c + c^2 + \cdots + c^{m-n-1})|x_1 - x_0| = c^n|x_1 - x_0| \sum_{j=0}^{m-n-1} c^j$$

Thus we have that,

$$|x_m - x_n| \leq c^n(1 + c + c^2 + \cdots + c^{m-n-1})|x_1 - x_0|$$

Note however that from Theorem (1.7.2) $\lim_{n\to\infty} x^m = r$ where r denotes the fixed point of g. So,

$$|r - x_n| = \lim_{m\to\infty} |x_m - x_n| \leq c^n|x_1 - x_0| \sum_{j=0}^{\infty} c^j$$

However $\sum_{j=0}^{\infty} c^j$ is a geometric series with $c < 1$ so we can easily sum it:

$$\sum_{j=0}^{\infty} c^j = \frac{1}{1-c}$$

Thus

$$|r - x_n| \leq \frac{c^n}{1-c}|x_1 - x_0|$$

Note that for both of the inequalities (1.6) and (1.7) the rate of convergence depends on the value of c. Thus the smaller the c the faster the convergence.

This is what went wrong with our previous example. The derivative of $g(x) = x^2 - 6 + x$ is $g^{'}(x) = 2x + 1$ which is definitely not less than 1 in the interval $[2, 10]$. In fact it is greater than 1. It is not surprising therefore that the method diverged from, instead of converged to, the fixed point.

Similarly we can obtain an a-posterriori type estimate and we therefore list it here without proof.

Corollary 1.8.2. *Suppose g as in Theorem (1.7.2) above. Then the error of our approximating the fixed point r of g can be estimated by*

$$|x_n - r| \leq \frac{c}{1-c}|x_n - x_{n-1}|$$

Let us now look at another example where we pay some more attention to the derivative of $g(x)$.
Example:
Solve the equation $f(x) = x^3 - 2x - 5 = 0$ in the interval $[2, 3]$. Let us first rewrite this in the form of a fixed point in two different ways:

$$\begin{aligned} x = g_1(x) &= \frac{x^3 - 5}{2} \\ x = g_2(x) &= (2x + 5)^{1/3} \end{aligned}$$

Solution:
The solution can now be easily found by applying the pseudocode,

n	0	1	2	3	4
$g_1(x)$	1.50000	$-.8125000$	-2.7681885	-13.1061307	-1128.1243667
$g_2(x)$	2.0800838	2.0923507	2.0942170	2.0945007	2.0945438

Thus clearly from the above $g_2(x)$ was a good function to use in order to approximate the root while $g_1(x)$ is not! The reason is really the theorem we just proved. Note that the respective derivatives are

$$\begin{aligned} g_1'(x) &= \frac{3}{2}x^2 > 1 \quad \text{in the interval } [2,3] \\ g_2'(x) &= \frac{2}{3\sqrt{(2x+5)^2}} << 1 \quad \text{in the interval } [2,3] \end{aligned}$$

which explains why $g_2(x)$ produced a solution and not $g_1(x)$ using the fixed point iteration.

1.9 Convergence and higher order methods

The question of convergence is of great importance for any numerical method which we will encounter. As such we need to devote more time in understanding how to find the convergence rates of some of the schemes which we have seen so far.

We start by defining a very important quantity: the rate of convergence. In some texts this is also called the speed of convergence.

Definition 1.9.1. *We say that* $\{x_n\}_{n=0}^{\infty}$ **converges to** x^* **with order** α *if there exist a constant* $\mu \in [0,1]$ *such that*

$$\lim_{n\to\infty} \frac{|x_{n+1} - x^*|}{|x_n - x^*|^\alpha} = \mu$$

We call μ *the rate of convergence and for* $\alpha = 1$ *must be a finite number in* $[0,1]$. *For other* α *values it is sufficient that* $\mu < \infty$ *for convergence.*

Clearly if $\mu > 1$ then the sequence $\{x_n\}_{n=0}^{\infty}$ diversges. Futhermore, we distinguise convergence as follows:
If $\alpha = 1$ then the convergence is *linear*.
If $\alpha = 2$ then the convergence is *quadratic*, ... etc.
If the sequence is linearly convergent and $\mu = 0$ then in fact the convergence is *superlinear*.
If instead $\mu = 1$ then the convergence is *sublinear*.
Finally if the sequence is linearly convergent and furthermore

$$\lim_{n\to\infty} \frac{|x_{n+1} - x_n|}{|x_n - x_{n-1}|} = 1$$

then the sequence $\{x_n\}_{n=0}^{\infty}$ *converges logarithmically* to x^*. Alternatively you could plot $\lim_{n\to\infty} |x_{n+1} - x^*|$. If the resulting plot is a straight line then the sequence converges logaritmically.

An example of a linearly converging sequence for instance would be

$$\{1, 1/2, 1/4, 1/8, 1/16, \dots\}$$

An example of a superlinearly converging sequence would be

$$\{1/2, 1/4, 1/16, 1/256, \dots .\}$$

Check the sequence above yourself to see why it is superlinearly convergent. In fact you could compute the order of convergence, α, for this sequence to be 2 which shows that it converges quadratically. As a final example we also provide a sequence which is logarithmically convergent, $\{1/\log(n)\}_{n=1}^{\infty}$.

Now we are in position to find out the order of convergence for some of the methods which we have seen so far. We start with the fixed point iteration.

Theorem 1.9.2. *The fixed point iteration method*

$$x_n = g(x_{n-1})$$

starting with an arbitrary x_0 converges linearly to the unique fixed point x under the assumption

$$0 \neq |g'(x)| \leq c < 1$$

with asymptotic error constant $|g'(x)|$.

Proof: Since the fixed point is x then $g(x) = x$. Then the following holds:

$$x - x_n = g(x) - g(x_{n-1}) = g'(\xi)(x - x_{n-1})$$

by the mean value theorem for some $\xi \in (x, x_{n-1})$. If we rewrite the above and take the limit as $n \to \infty$

$$\lim_{n\to\infty} \frac{|x - x_n|}{|x - x_{n-1}|} = \lim_{n\to\infty} |g'(\xi)|$$

Remember however that for the mean value theorem $\xi \in (x, x_{n-1})$. Since $x_n \to x$ then also $\xi \to x$. As a result,

$$\lim_{n\to\infty} |g'(\xi)| = |g'(x)|$$

And therefore the result,

$$\lim_{n\to\infty} \frac{|x - x_n|}{|x - x_{n-1}|} = |g'(x)|$$

Note that the above indicates the convergence rate for the fixed point iteration to be linear (since here $\alpha = 1$) with asymptotic constant $|g'(x)|$. In fact, as we will see, many of the methods which we will learn converge linearly.

Once again, we must emphasize our earlier remark, based on the Corollary, which states that for the fixed point iteration convergence is guaranteed if $|g'(x)| \leq c < 1$. In particular the smaller the c is the faster the convergence!

1.10 Improving Convergence

Although we would rather have quadratic convergence it is actually not that easy to device methods which find the root of $f(x) = 0$ that fast. Newton's method is the only one which we have seen so far which could achieve quadratic convergence under certain conditions. Most of the other methods which we have seen converge linearly.

However even among linearly convergent methods some converge faster than others. In fact among those linear methods which we have seen the smaller the *asymptotic error constant* is the faster the convergence (but still linear). In fact we have already introduced the notion of *super-linear* convergence. We provide below a formal definition.

Definition 1.10.1. *A sequence $\{x_n\}$ is said to converge super-linearly to x if*

$$\lim_{n\to\infty} \frac{x_{n+1} - x}{x_n - x} = 0$$

We will now look at a technique which allows us to improve the convergence of a given sequence.

1.11 Aitken's Method

Aitken's method is a very general method for improving convergence of any sequence. With that perspective we put into use this method in order to facilitate our task of obtaining the root of a function faster than before. Note also that Aitken's method is applied on an already existing sequence in order to speed it up. But first the derivation of the method.

One way to derivate Aitken's is simply by noting the following equivalence in the case of sufficiently large n:

$$\frac{x_{n+1} - r}{x_n - r} \equiv \frac{x_{n+2} - r}{x_{n+1} - r}$$

where r, as usual, denotes the root or solution to $f(x) = 0$. The above is clearly true when n is large enough. Rewriting we obtain,

$$(x_{n+1} - r)^2 \equiv (x_{n+2} - r)(x_n - r)$$

Expanding both sides of the above we obtain,

$$r(x_{n+2} + x_n - 2x_{n+1}) \equiv x_{n+2}x_n - x_{n+1}^2$$

Thus solving for r we have,

$$r \equiv \frac{x_{n+2}x_n - x_{n+1}^2}{x_{n+2} - 2x_{n+1} + x_n} \tag{1.8}$$

Working on the numerator only by adding and subtracting the same quantities we get,

$$\begin{aligned} x_{n+2}x_n - x_{n+1}^2 &= x_{n+2}x_n - x_{n+1}^2 + 2x_nx_{n+1} - 2x_nx_{n+1} + x_n^2 - x_n^2 \\ &= x_n(x_n - 2x_{n+1} + x_{n+2}) - (x_n^2 - 2x_nx_{n+1} + x_{n+1}^2) \end{aligned}$$

Substituting this in (1.8) we obtain the iteration for Aitken's method:

$$\hat{x}_n = x_n - \frac{(x_{n+1} - x_n)^2}{x_{n+2} - 2x_{n+1} + x_n}$$

The pseudo-code for Aitken's scheme is provided below.

1. Begin by reading the first 3 elements $x_i = x_0, x_j = x_1$ and $x_k = x_2$ of the given sequence which we want to improve the convergence of.

2. Obtain the next element for the new improved sequence $\{\hat{x}_n\}$ from Aitken's method,

$$\hat{x}_n = x_i - \frac{(x_j - x_i)^2}{x_k - 2x_j + x_i}$$

3. If no more elements in $\{x_n\}$ then Stop. Otherwise read the next three elements from the sequence $x_i = x_j, x_j = x_k$ and $x_k = x_{k+1}$ and calculate again from step 2 above.

This new sequence $\hat{x}_n$ converges faster to the root r than the original sequence x_n would. Note that Aitken's method can be used on an existing sequence for improvement. In other words if the given sequence was supposed to find a root then Aitken's method will use that original sequence and improve it thus obtaining the answer faster. However, keep in mind that Aitken's scheme can be used on any sequence at all (not just root finding ones).
Example: Using the following sequence:

n	1	2	3	4	5	6	7
x_n	1.00000	1.25992	1.31229	1.32235	1.32426	1.32463	1.32470

improve its convergence using Aitken's method.
Solution:
Note that here we are not given a function or any clue at all where this sequence originated. We are just asked to improve its convergence.

n	1	2	3	4	5
$\hat{x}_n$	1.32550	1.32474	1.32470	1.32471	1.32471

The following theorem helps in understanding what is the improvement in convergence for Aitken's method.

Theorem 1.11.1. *Suppose $\{x_n\}_{n=0}^{\infty}$ a sequence which converges linearly to its limit r with asymptotic constant less than 1 and $x_n - r \neq 0$ for all $n \geq 0$. Then the new sequence $\{\hat{x}_n\}_{n=0}^{\infty}$ as defined above converges also to r with 0 asymptotic constant,*

$$\lim_{n\to\infty} \frac{\hat{x}_n - r}{x_n - r} = 0$$

1.12 Steffensen's Method

Although we just presented a way to improve convergence of a given sequence (and therefore a given method) we will now present a method which can be even faster. Steffensen's method which we will present now is a modification of Aitken's method but can not be applied to just any sequence. Steffensen's method has to be applied to a sequence solving a fixed point problem.

Steffensen's scheme is outlined below.

1. Obtain a starting guess x_0 as well as two more iterates through the usual fixed point iteration: $x_1 = g(x_0)$ and $x_2 = g(x_1)$.

2. Use Aitken's method now to obtain x^* from

$$x^* = x_0 - \frac{(x_1 - x_0)^2}{x_2 - 2x_1 + x_0}$$

3. If required tolerance has been reached or maximum number of iterates has been exceeded then stop with appropriate message.

4. Otherwise let $x_0 = x^*, x_1 = g(x_0)$ and $x_2 = g(x_1)$ and continue calculating from step 2.

Example: Solve the fixed point problem $x = \frac{2-e^x+x^2}{3}$ in the interval $[0, 1]$ with accuracy 10^{-4} using fixed point iteration, Aitken's and Steffensen's method.
Solution:
We display below all three methods:
Fixed Point Iteration:

n	1	2	3	4	5	6	7	8	9
x_n	1.00	.093906	.303454	.245852	.260578	.256740	.257735	.257477	.257544

Aitken's

n	1	2	3	4	5	6	7
$\hat{x}_n$	.264095	.258272	.257579	.257534	.257531	.257530	.257530

Steffensen's

n	1	2	3
x_n^*	.264095	.257531	.257530

As mentioned in the beginning of this section Steffensen's is fast! Faster than Aitken's method in any case. But Aitken's scheme was super-linear and as such we would expect Steffensen's method to be quadratic, like Newton-Raphson is. In fact the following theorem gives us the conditions for exactly that:

Theorem 1.12.1. *Suppose that the fixed point problem $g(x) = x$ has solution r and $g'(r) \leq 1$. If there exist a $\delta > 0$ such that $g \in C^3[r - \delta, r + \delta]$ then Steffensen's method gives quadratic convergence for any $x_0 \in [x + \delta, x - \delta]$.*