Nonlinear Equations: Fixed Point

Math 131: Numerical Analysis

J.C. Meza

2/15/24

Fixed Point

Methods for root finding are hard to analyze, especially if we include derivatives of f. Another approach is to reformulate the problem to allow for easier analysis using what is known as a fixed point formulation. This approach has a long history and is also known as functional iteration, Picard iteration, and successive substitution.

Note:

We will do everything in $\mathbb{R}^1$ but the ideas hold in higher dimensions with appropriate generalizations.

First let's start with a definition.

Fixed Points

Definition: The point $x^* \in \mathbb{R}$ is said to be a ***fixed point*** of $g : \mathbb{R} \to \mathbb{R}$ if $g(x^*) = x^*$.

Using the definition it is easy to see that finding fixed points is equivalent to finding roots of an equation.

For example, let's define a function $g(x)$ such that

$$g(x) = x - f(x)$$

Then x^* is a zero of $f(x)$ if and only if x^* is a fixed point of $g(x)$, i.e.

$$f(x^*) = 0 \iff g(x^*) = x^*$$

Equivalence of zeros and fixed points

By definition:

$$g(x) = x - f(x).$$

If x^* is a zero of $f(x)$ then clearly $g(x^*) = x^*$.

On the other hand, if $g(x^*) = x^*$ then

$$f(x^*) = x^* - g(x^*) = x^* - x^* = 0$$

Note

Once we know how to find fixed points, we can find zeros of a function.

Equivalence of Fixed Points and Zeros

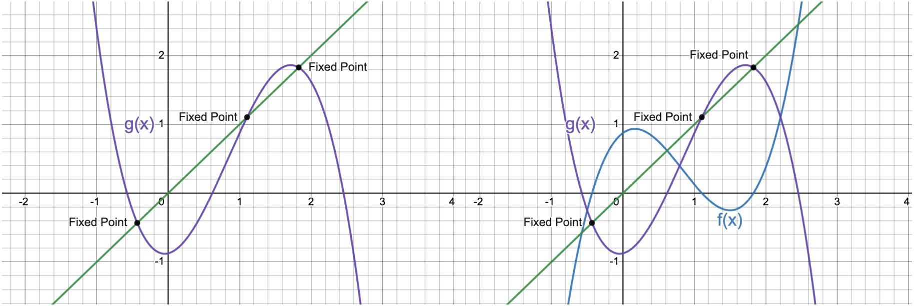


Figure 1: Fixed Point is where g(x) meets y=x. Zeros align with fixed points

Idea

- The general idea is that if we have an approximation. x_0, to the fixed point x^*, then if we take $x_1 = g(x_0)$, our hope is that x_1 will be an even better approximation.
- There are many choices we can make for $g(x)$ for any given $f(x)$. And not too surprisingly, the choice of $g(x)$ will be extremely important in determining whether we produce a good algorithm or not.
- In the rest of the section, we will look into some of the details of this approach and what conditions are needed to ensure that the iteration converges.

Algorithm

```python
# Example: Fixed point iteration
maxiter = 50
x = x0
for k in range(1,maxiter):
    x = g(x)
    if (abs(g(x)-x) < 1.e-15):
        break
xsol = x
```

Example

Let's suppose that the sqrt key on your calculator is broken and you need to compute $\sqrt{3}$. This is equivalent to finding a solution of the function $f(x) = x^2 - 3 = 0$.

First attempt:

$g(x) = 3/x, x_0 = 1$

Using the algorithm above we get the following sequence of iterates:

$$
\begin{aligned}
x_0 &= 1 \\
x_1 &= 3/x_0 = 3 \\
x_2 &= 3/x_1 = 1 \\
x_3 &= 3/x_2 = 3 \\
&\vdots
\end{aligned}
$$

Apparently, this is not a good choice!

Second attempt:

$g(x) = \frac{1}{2}(x + 3/x), x_0 = 1$

This time the sequence of iterates is:

$$
\begin{aligned}
x_0 &= 1 \\
x_1 &= \frac{1}{2}(x_0 + 3/x_0) = 2 \\
x_2 &= \frac{1}{2}(x_1 + 3/x_1) = 1.75 \\
x_3 &= \frac{1}{2}(x_2 + 3/x_2) = 1.73214 \\
x_4 &= \frac{1}{2}(x_3 + 3/x_3) = 1.7320508
\end{aligned}
$$

Note that after only 4 iterations, our solution appears to be a great approximation to $\sqrt{3}$. Later on, we'll find out why this choice of $g(x)$ is a good choice.

In class exercise:

Compute $\sqrt[3]{5}$ using the same procedure as above with $g(x) = \frac{1}{3}(2x + 5/x^2),\ x_0 = 1.$

$$\begin{aligned}
x_0 &= 1 \\
x_1 &= \frac{1}{3}(2x_0 + 5/x_0^2) = \\
x_2 &= \\
x_3 &= \\
x_4 &= \\
&\vdots
\end{aligned}$$

Existence and Uniqueness of Fixed Points

At this point, there are several questions that arise naturally:

1. Is there a fixed point x^* in the interval $[a, b]$?
2. If yes, is it unique?
3. Does a given fixed point iteration generate a sequence of iterates $\{x_k\} \to x^*$?
4. And if yes, how fast will the iterates converge to the fixed point?

We'll take each of these points in turn starting with (1) and (2) using the following theorem.

Theorem: Fixed Point Existence/Uniqueness

If $g \in C[a,b]$ and $g(x) \in [a,b]\ \forall x \in [a,b]$, then there exists a fixed point $x^* \in [a,b]$.

If in addition, $g'(x)$ ***exists*** on (a,b) and there exists a positive constant $k < 1$ with $|g'(x)| \leq k \quad \forall x \in (a,b)$ then there exists ***exactly 1***. fixed point in $[a,b]$.

The first part of the theorem states the condition for the existence of a fixed point. The second part states the conditions necessary for uniqueness. Let's take them one at a time.

Existence

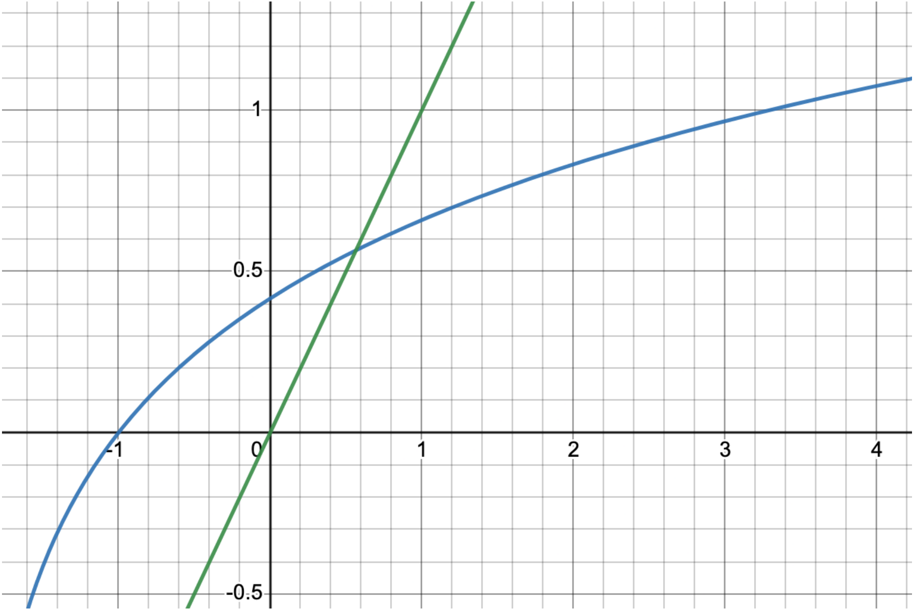
There are two cases to consider. The first is if the fixed point is one of the endpoints of the interval $[a, b]$. The second case, is if the fixed point is in the interior.

If $g(a) = a$ or $g(b) = b$, then clearly the fixed point $x^* = a$ or b, so we're done.

If not, then if the fixed point exists it must lie in the interior, i.e. $g(a) \neq a$ and $g(b) \neq b$. By assumption $g(x)$ maps the interval $[a, b]$ onto itself, so we can deduce that

$$g(a) > a \quad \text{and} \quad g(b) < b.$$

Visually



Existence Proof (cont.)

Let's now consider the function

$$h(x) = g(x) - x$$

First note that $h(x)$ is continuous on $[a, b]$ since $g(x)$ is continuous. Next, it is easy to see that it also satisfies the conditions:

$$h(a) = g(a) - a > 0 \quad \text{and} \quad h(b) = g(b) - b < 0$$

By the Intermediate Value Theorem, there exists a point $x^* \in (a, b)$ such that $h(x^*) = 0$. But using the definition of $h(x)$ that means that

$$\begin{aligned} h(x^*) = 0 &= g(x^*) - x^* \\ &\implies g(x^*) = x^* \end{aligned}$$

So x^* must be a fixed point in $[a, b]$.

Uniqueness

To show that a unique fixed point exists we will have need of the second condition

$$|g'(x)| \leq k < 1$$

The proof will be by contradiction. Let's assume that we have two fixed points x^*, y^* and that $x^* \neq y^*$. Using the Mean Value Theorem we can say that:

$$\frac{g(x^*) - g(y^*)}{x^* - y^*} = g'(\xi) \quad \xi \in [x^*, y^*] \subset [a, b].$$

Uniqueness Proof (cont.)

Let's now consider $|x^* - y^*|$:

$$|x^* - y^*| = |g(x^*) - g(y^*)| = |g'(\xi)| \cdot |x^* - y^*| < k \cdot |x^* - y^*|$$

The first equation is true by the definition of a fixed point, and the second from the equation above derived from the MVT. The final inequality is due to the assumption on the bound of the derivative of g. But since the constant $k < 0$, this reduces to:

$$|x^* - y^*| < |x^* - y^*|$$

which is a contradiction.

Therefore, x^* must be unique. ■

More on convergence of sequences

- The question of convergence is not simply a matter of deciding when one should stop an iterative method.
- When one has a choice of different algorithms to pick from, it would make sense to choose the one that is fastest, where fastest can be loosely defined to be the one that is likely to take the fewest number of iterations.
- Towards that end, let's discuss convergence in a bit more detail.

Definition: Rate of Convergence

We say that the ***rate of convergence*** of x_k to x^* is of ***order*** $p \geq 1$ if

$$\lim_{k \to \infty} \frac{|x_{k+1} - x^*|}{|x_k - x^*|^p} < \infty$$

Remark

If $p = 2$, we say the rate of convergence is ***quadratic***. All other things being equal, the larger the value of p, the faster an algorithm will converge to a solution.

Rate of convergence (cont.)

For the special case of $p = 1$ we ask instead that

$$\lim_{k\to\infty} \frac{|x_{k+1} - x^*|}{|x_k - x^*|} < 1$$

This is known as a ***linear*** rate of convergence.

In addition, if for some sequence $\{c_k\} \to 0$ we have

$$|x_{k+1} - x^*| \leq c_k \, |x_k - x^*|$$

then the sequence $\{x_k\}$ is said to converge ***superlinearly*** to x^*.

Summary

A quadratic rate of convergence is better than a superlinear rate of convergence, which is better than linear.

Examples.

1. $x_k = a^k, \quad 0 < a < 1$
2. $x_k = a^{2^k}, \quad 0 < a < 1$
3. $x_k = \left(\frac{1}{k}\right)^k$

Solutions:

(1) linear, (2) quadratic, (3) superlinear.