Nonlinear Equations: Fixed Point Iteration Convergence

Math 131: Numerical Analysis

J.C. Meza

2/20/24

Last lecture

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?

Fixed Point Iteration Convergence (Part 1)

Now that we've answered questions (1) and (2) on the existence and uniqueness of fixed points, we will turn our attention to the last two questions dealing with the convergence of the fixed point iteration itself.

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?

Recall: Fixed Point Iteration

```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
```

In order to better understand the possible cases, let's take a look at a couple of pictures.

Visually - Convergence

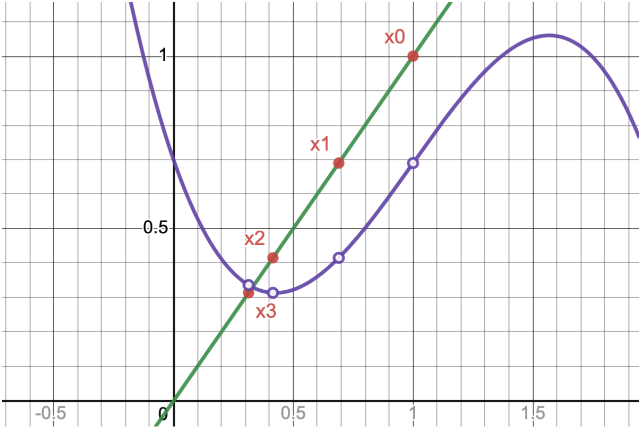

Figure 1: Fixed Point Iteration Converges

Visually - Divergence

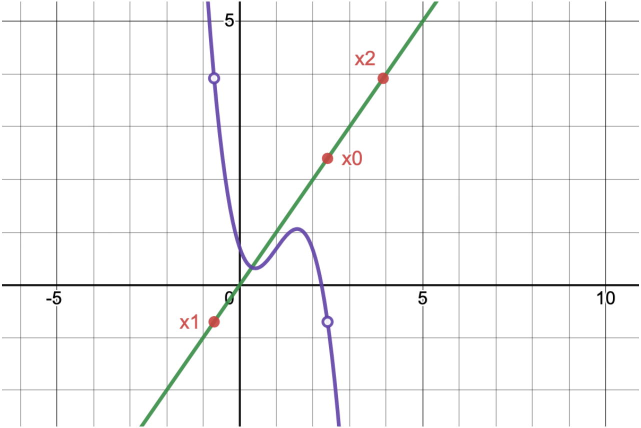

Figure 2: Fixed Point Iteration can diverge

Discussion

- If the slope of the function $g(x)$ is too large, then the fixed point iteration might not converge.
- Another way to think about this is that the derivative must be bounded in some way for the fixed point iteration to converge.
- Your textbook proves convergence by making an assumption on the derivative of $g(x)$ and specifically that there is a constant $k < 1$ such that:

$$|g'(x)| \leq k \quad \forall x \in [a, b]$$

We will use two other concepts that will prove useful in later lectures. It also has the additional benefit of not requiring the assumption that the derivative of $g(x)$ exists.

Lipschitz Continuity

Definition. A function $g : \mathbb{R} \to \mathbb{R}$ is said to be ***Lipschitz continuous*** if

$$|g(x) - g(y)| \leq L\ |x - y| \quad \forall x \in [a, b].$$

L is called the ***Lipschitz constant.***

Important

A function that is Lipschitz continuous is bounded in how much it can change. However, the definition only requires the function be bounded ***on an interval*** and not over the entire domain of the function.

Relation to derivative

As it turns out there is a close relation of the Lipschitz constant to the derivative. Recall that by the MVT

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

Taking absolute values of both sides we get:

$$\left|g(x) - g(y)\right| = \left|g'(\xi)\right| \cdot \left|x - y\right| \quad \xi \in (x, y) \subset [a, b].$$

Proof (cont.)

Now, if we let

$$L = \max_{\xi \in [a,b]} \left| g'(\xi) \right|,$$

it follows that:

$$\left| g(x) - g(y) \right| \leq L \cdot \left| x - y \right| \quad x \in [a,b].$$

Intuitively this makes sense as bounding the derivative is similar to bounding the change in the function values.

Contraction Mapping

Definition. We say that $g : \mathbb{R} \to \mathbb{R}$ is a ***contraction mapping*** if

$$|g(x) - g(y)| \leq L \; |x - y| \quad \forall x \in [a, b].$$

with $L < 1$.

Note

Notice the close similarity to Lipschitz continuity, but with the important distinction that $L < 1$.

Example: Contraction Mapping

Consider

$$f(x) = e^x - 2x - 1 \quad \text{on } [1, 2]$$
$$g(x) = \ln(2x + 1)$$

Show that $g(x)$ is a contraction mapping. Note: $x^* \approx 1.256$

Solution Outline

1. Compute g'
2. Note that g' is monotonically decreasing on given interval
3. Show bound on g', and in particular that the bound is < 1.

We're now ready to show our first convergence theorem for fixed point iterations.

Fixed Point Convergence Theorem

Theorem. Suppose $g : \mathbb{R} \to \mathbb{R}$, with $g(x) \in [a, b] \quad \forall x \in [a, b]$ be a continuous contraction mapping. Then there exists a unique fixed point x^* with $g(x^*) = x^*$ and the fixed point iteration converges to x^* for any starting point $x_0 \in [a, b]$.

Proof:

Existence and uniqueness of a fixed point was proven earlier.

To prove convergence of the fixed point iteration we need to show that $\{x_k\} \to x^*$ for the fixed point iteration:

$$x_{k+1} = g(x_k)$$

Proof (cont.)

Let's consider: $x_k - x^* = g(x_{k-1}) - g(x^*)$

Taking absolute values and using the fact that g is a contraction mapping we can write

$$\left|x_k - x^*\right| = \left|g(x_{k-1}) - g(x^*)\right| \leq L \left|x_{k-1} - x^*\right|.$$

Now we apply this inductively:

$$\begin{aligned}\left|x_k - x^*\right| &\leq L^2 \left|x_{k-2} - x^*\right|. \\ &\leq \cdots \\ &\leq L^k \left|x_0 - x^*\right|\end{aligned}$$

Proof (cont.)

But since $L < 1$ we know that $L^k \to 0, k \to \infty$ and therefore

$$\left\|x_k - x^*\right\| \to 0, k \to \infty$$

as we set out to show. ■

Remarks

- As it turns out, while a contraction mapping is a useful tool, it is often hard to verify in a real-world application.
- Instead what is usually assumed is that either the function is Lipschitz continuous in a neighborhood of the fixed point x^*, or that the function has a bounded derivative at the root.
- Another thing to keep in mind is that it might be unrealistic to assume a contraction mapping property for all $x \in [a, b]$. An alternative is to assume that we have some property that holds at or near the solution. In this case, we have a different type of convergence.

Local Convergence

Definition. The iterative process $x_{k+1} = g(x_k)$ is said to be ***locally convergent*** to the fixed point x^* if there exists $\delta > 0$ such that $x_k \to x^*$ for any x_0 satisfying $|x^* - x_0| \le \delta$.

> Tip
>
> The simplest interpretation of this definition is to say, if we start ***close enough*** to a fixed point, then the method will converge to it.

What defines close enough, will depend not only on the initial point, but also on the function we use.

With this background, we can state the following theorem.

Local Convergence Theorem

Theorem. Suppose $g : \mathbb{R} \to \mathbb{R}$ is continuously differentiable and:

1. $x^* = g(x^*)$
2. $|g'(x^*)| < 1$

Then the iterative process $x_{k+1} = g(x_k)$ is ***locally convergent*** to x^*.

Remark: Note the difference here that we are only asking for a bound on g' at a single point, namely the fixed point, x^*.

The proof follows much the same as before, but now we have to also show that $\left|g'(x)\right| < 1$ for all x in a neighborhood of x^*. Since we are assuming that g is continuously differentiable, this is fairly easy to show.

Fixed Point Iteration Convergence (Part 2)

Theorem. Suppose $g : \mathbb{R} \to \mathbb{R}$ is twice continuously differentiable and the iterative process $x_{k+1} = g(x_k)$ is locally convergent.

Then the convergence rate is ***linear*** if

$$g'(x^*) \neq 0$$

and the convergence rate is at least ***quadratic*** if

$$g'(x^*) = 0.$$

Proof - Linear case

We've already shown that the fixed point iteration converges.

To show that we have linear convergence, we can either

1. show that we can bound the derivative in a neighborhood of the fixed point or
2. we can use the Lipschitz continuity condition, to show that the error at any iteration is bounded by $L < 1$, which gives us a linear rate of convergence.

Proof - Quadratic case

To show that the convergence rate is at least quadratic, let's first expand $g(x)$ in a Taylor polynomial about the point x^*.

$$g(x) = g(x^*) + g'(x^*)(x - x^*) + \frac{g''(\xi)}{2}(x - x^*)^2. \tag{1}$$

By assumption we know that:

$$g(x^*) = x^* \text{ and } g'(x^*) = 0.$$

Proof (cont.)

Substituting into (Equation 1), we get:

$$g(x) = x^* + 0 + \frac{g''(\xi)}{2}(x - x^*)^2.$$

Consider $x = x_k$:

$$g(x_k) = x^* + \frac{g''(\xi_k)}{2}(x_k - x^*)^2,$$

with ξ_k between x_k and x^*.

Proof (cont.)

Using the definition of the fixed point iteration: $x_{k+1} = g(x_k)$, we have

$$x_{k+1} = x^* + \frac{g''(\xi_k)}{2}(x_k - x^*)^2,$$

Rearranging we have:

$$\frac{x_{k+1} - x^*}{(x_k - x^*)^2} = \frac{g''(\xi_k)}{2},$$

Proof (cont.)

Since the fixed point iteration is locally convergent this means that

$$\{x_k\} \to x^* \implies \{\xi_k\} \to x^*$$

since ξ_k is between x_k and x^*.

Therefore

$$\frac{x_{k+1} - x^*}{(x_k - x^*)^2} = \frac{g''(x^*)}{2} = C < \infty,$$

which proves that the iterates are converging q-quadratically.

Designing good g(x) functions

Important

This theorem shows us that if we want quadratic convergence, we should look for fixed point methods with the property that $g'(x^*) = 0$.

Let's consider the general form

$$g(x) = x - \phi(x)f(x) \tag{2}$$

where $\phi(x)$ is differentiable and to be chosen later.

Taking derivatives we have:

$$g'(x) = 1 - \phi'(x)f(x) - f'(x)\phi(x).$$

Letting $x = x^*$, such that $f(x^*) = 0$, we get:

$$g'(x^*) = 1 - f'(x^*)\phi(x^*).$$

Conditions for Quadratic Convergence

By our theorem, if we want quadratic convergence we need to have $g'(x^*) = 0$, hence

$$0 = 1 - f'(x^*)\phi(x^*),$$

and solving for ϕ we get the result:

$$\phi(x^*) = \frac{1}{f'(x^*)}.$$

Substituting back into our general form (Equation 2) we have:

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

But that's just Newton's method!

As a result, an immediate corollary is that Newton's method is quadratically convergent.

Example

Recall our earlier example where we wanted to find the root of $x^2 - 3$, that led us to try different fixed point iterations.

1. $g(x) = 3/x, \implies g'(x) = -3/x^2.$

$$g'(x^*) = g'(\sqrt{3}) = -1 \quad \implies \quad \text{No convergence}$$

2. $g(x) = \frac{1}{2}(x + 3/x), \implies g'(x) = \frac{1}{2}(1 - 3/x^2).$

$$g'(x^*) = g'(\sqrt{3}) = 0 \quad \implies \quad \text{q-quadratic convergence}$$

3. $g(x) = x - \frac{x^2-3}{7}, \implies g'(x) = 1 - \frac{1}{7}(2x).$

$$g'(x^*) = g'(\sqrt{3}) = 0.505 \quad \implies \quad \text{Linear convergence}$$

Estimating number of iterations

- Sometimes it is useful to estimate the number of iterations required to achieve a certain reduction in the error.
- We were able to do this in the case of the bisection method as each iteration reduced the interval in half, so therefore it was easy to compute the number of iterations needed to reduce the error by a certain factor.
- We can do something similar for other methods, although it can be a bit trickier.

Convergence (cont.)

Suppose that we know that we have q-linear convergence with a constant λ, i.e.

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

One way to approach this is to consider our proof of convergence for the fixed point iteration. We were able to show that

$$|x_k - x^*| \leq L^k \, |x_0 - x^*|.$$

Iterations to Convergence

Suppose we would like to reduce the error by a factor of 10 after k iterations. Setting $\lambda = L$, this means that we want to have:

$$\lambda^k = 10^{-1}$$

Taking logs of both sides we have:

$$k \; \log_{10} \lambda = -1$$

or equivalently

$$k = -\frac{1}{\log_{10} \lambda}$$

The value of k can be interpreted to mean that it approximates the number of iterations required to reduce the error by a factor of 10.

Example

Suppose that we knew that a certain fixed point iteration g had a Lipschitz constant of $L = 2/3$. Then setting $\lambda = L$.

$$k = -\frac{1}{\log_{10}(2/3)} \approx \frac{1}{0.176} = 5.68$$

which we can interpret to mean that it would take approximately $k = 6$ iterations to reduce the error by a factor of 10.

Summary Comparison of nonlinear equation methods

Method	Assumptions	Advantages	Disadvantages
Bisection	f is continuous; 2 starting points where function has opposite sign	Robust; easy to implement	Linear convergence
Newton	f is continuously differentiable	Quadratic convergence	Need derivatives
Secant	f is continuous	No derivatives; superlinear convergence	Cancellation error possible
Fixed Point	Need to have a good $g(x)$	Easy to implement	Linear convergence