Nonlinear Equations: Bisection and Fixed-Point Iteration

1 Solving Nonlinear Equations

We consider the nonlinear equation

$$f(x) = 0.$$

In general, a nonlinear equation need not have a closed-form solution. Even when an exact solution exists, it may be difficult or impractical to compute. We therefore seek numerical methods that construct approximations to a root.

A basic difficulty is that nonlinear equations may have multiple roots, no roots, or roots whose location is not obvious from the formula alone. A graph of f can often provide useful preliminary information.

2 The Bisection Method

Suppose that f is continuous on an interval $[a, b]$ and

$$f(a)f(b) < 0.$$

By the Intermediate Value Theorem, there is at least one point $x^* \in (a, b)$ such that

$$f(x^*) = 0.$$

The bisection method repeatedly cuts the interval in half. Let

$$c = \frac{a+b}{2}.$$

We evaluate $f(c)$. If $f(c) = 0$, then we have found a root. Otherwise, we keep the half-interval on which the sign change occurs:

$$\begin{cases} [a, c], & f(a)f(c) < 0, \\ [c, b], & f(c)f(b) < 0. \end{cases}$$

Repeating this process produces nested intervals containing a root.

Algorithm

Starting with $[a_0, b_0]$ such that $f(a_0)f(b_0) < 0$, for $n = 0, 1, 2, \ldots$:

1. Set
$$c_n = \frac{a_n + b_n}{2}.$$
2. If $f(c_n) = 0$, stop.

3. If $f(a_n)f(c_n) < 0$, set

$$a_{n+1} = a_n, \qquad b_{n+1} = c_n.$$

Otherwise set

$$a_{n+1} = c_n, \qquad b_{n+1} = b_n.$$

After each iteration, the length of the interval is divided by two. Hence

$$b_n - a_n = \frac{b_0 - a_0}{2^n}.$$

If x^* denotes a root in $[a_n, b_n]$ and c_n is the midpoint, then

$$|x^* - c_n| \le \frac{b_n - a_n}{2} = \frac{b_0 - a_0}{2^{n+1}}.$$

Thus, to guarantee an error no larger than a prescribed tolerance ε, it is sufficient to choose n so that

$$\frac{b_0 - a_0}{2^{n+1}} \le \varepsilon.$$

Stopping criteria

In practice, several stopping criteria are possible:

$$|f(c_n)| < \varepsilon, \qquad |b_n - a_n| < \varepsilon,$$

or a related estimate based on the midpoint error.

A small residual $|f(c_n)|$ does not always imply a small error in c_n. The geometry of the function near the root matters. The interval-width criterion has the advantage of giving a direct bound on the error in the independent variable.

Examples where bisection can fail or mislead

A major advantage of the bisection method is that, once a sign-changing bracket has been found, it requires only that f be continuous. In particular, no derivatives are needed. This makes its assumptions weaker than those of several other root-finding methods we will discuss. The bracketing condition itself, however, can be restrictive, and a valid bracket does not necessarily identify a particular root when several roots are present.

Example 1: a root that cannot be bracketed by a sign change. Consider

$$f(x) = x^2.$$

The function has a root at $x = 0$, but $f(x) \ge 0$ for every x. Hence there is no interval $[a, b]$ containing the root for which

$$f(a)f(b) < 0.$$

Thus the usual bisection hypothesis fails even though a root exists. More generally, an even-multiplicity root need not produce a sign change.

Example 2: a valid bracket containing many roots. Let

$$f_N(x) = \cos\left(Nx + \frac{1}{10}\right), \qquad x \in [0, \pi].$$

This function has N roots in $[0, \pi]$. At the endpoints,

$$f_N(0) = \cos\left(\frac{1}{10}\right),$$

and

$$f_N(\pi) = \cos\left(N\pi + \frac{1}{10}\right) = (-1)^N \cos\left(\frac{1}{10}\right).$$

Therefore, when N is odd,

$$f_N(0)f_N(\pi) < 0,$$

so $[0, \pi]$ is a perfectly valid bracketing interval for bisection. Nevertheless, the interval contains N different roots. Bisection will converge to one of them, but the bracketing condition alone does not tell us which root will be selected. In particular, if a specific root is sought, bisection applied to the large interval $[0, \pi]$ may converge to a different root. One must first choose a bracket that isolates the desired root.

Bisection Method: Summary

Advantages.

- **Minimal assumptions.** The function f only needs to be continuous on the bracketing interval $[a, b]$, with

  $$f(a)f(b) < 0.$$

  In particular, no derivatives of f are required.

- **Guaranteed convergence.** Under these assumptions, bisection converges to a root contained in the interval.

- **Explicit error control.** Since the interval is halved at every iteration,

  $$b_n - a_n = \frac{b_0 - a_0}{2^n},$$

  and the midpoint approximation satisfies

  $$|c_n - x^*| \le \frac{b_0 - a_0}{2^{n+1}}.$$

  Thus, the number of iterations required to achieve a prescribed accuracy can be determined in advance.

- **Low cost per iteration.** After initialization, each iteration requires only one new function evaluation.

Limitations.

- **A bracketing interval is required.** We must first find a and b such that

$$f(a)f(b) < 0.$$

  This is not necessarily a disadvantage; it is part of what gives bisection its convergence guarantee.

- **Some roots cannot be bracketed by a sign change.** For example, $f(x) = x^2$ has a root at $x = 0$, but f does not change sign across the root.

- **A bracket need not isolate a unique root.** If several roots lie in $[a, b]$, bisection is guaranteed to find a root, but not necessarily the root we intended to find. The example

$$f_N(x) = \cos\left(Nx + \frac{1}{10}\right)$$

  illustrates this: for odd N, $[0, \pi]$ is a valid bracketing interval, yet it contains N roots.

- **Linear convergence.** The error is reduced by approximately a factor of two at each iteration. This is considerably slower than the convergence achievable by some methods that use additional information about f.

- **No direct multidimensional generalization.** The simple sign-change argument underlying scalar bisection does not extend directly to systems of nonlinear equations.

Key point. A sign-changing bracket guarantees that at least one root exists in the interval. It does not guarantee that the root is unique or that bisection will find a particular root of interest.

3 Fixed-Point Iteration

Another approach is to rewrite an equation in the form

$$x = g(x).$$

A solution x^* satisfying

$$x^* = g(x^*)$$

is called a *fixed point* of g.

This suggests the iteration

$$x_{n+1} = g(x_n), \qquad n = 0, 1, 2, \ldots,$$

starting from an initial guess x_0.

Geometrically, fixed points are intersections of the graphs

$$y = x \qquad \text{and} \qquad y = g(x).$$

The existence of an intersection does not by itself guarantee that the iteration converges to it. The behavior of g near the fixed point is essential.

4 Convergence of Fixed-Point Iteration

Suppose that

$$g : [a, b] \to [a, b]$$

is continuous on $[a, b]$ and continuously differentiable on (a, b), with

$$|g'(x)| < 1, \qquad x \in (a, b).$$

Under these assumptions, the fixed-point iteration

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

converges to the unique fixed point $x^* \in [a, b]$ for any initial value $x_0 \in [a, b]$.

If

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

on the interval, then the Mean Value Theorem gives

$$|g(x) - g(y)| \leq q|x - y|.$$

Therefore, if $x^* = g(x^*)$,

$$\begin{aligned} |x_{n+1} - x^*| &= |g(x_n) - g(x^*)| \\ &\leq q|x_n - x^*|, \end{aligned}$$

and hence

$$|x_n - x^*| \leq q^n |x_0 - x^*|.$$

Thus the fixed-point iteration converges linearly.

5 Example

Consider the fixed-point equation

$$x = \frac{1}{2}\sin x - \frac{1}{\pi}\sqrt{|x + 3|}.$$

Define

$$g(x) = \frac{1}{2}\sin x - \frac{1}{\pi}\sqrt{|x + 3|}.$$

A useful first step is to examine the graphs

$$y = x \qquad \text{and} \qquad y = \frac{1}{2}\sin x - \frac{1}{\pi}\sqrt{|x + 3|}.$$

Their intersection gives a fixed point of g and therefore a solution of the equation.

To justify fixed-point iteration on a chosen interval $[a, b]$, we then check two properties:

1. $g([a, b]) \subseteq [a, b]$;
2. $|g'(x)| \leq q < 1$ on $[a, b]$.

Once these are established, the iteration converges to the unique fixed point in the interval:

$$x_{n+1} = \frac{1}{2}\sin x_n - \frac{1}{\pi}\sqrt{|x_n + 3|}$$

for any starting value x_0 in that interval.

To be continued in class on Wednesday.