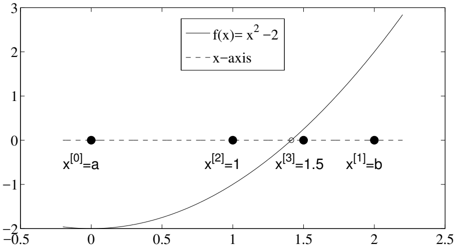
Solving nonlinear equations

1.1: The root-finding problem

MA385 – Numerical Analysis

September 2025



0. Outline

Linear equations are of the form:

$$\textit{find } x \textit{ such that } ax + b = 0$$

and are easy to solve. Some nonlinear problems are also easy to solve, e.g.,

$$\textit{find } x \textit{ such that } ax^2 + bx + c = 0.$$

Similarly, there are formulae for all cubic and quartic polynomial equations. But most equations do not have simple formulae for their solutions, so numerical methods are needed.

References

- Chap. 1 of Süli and Mayers (Introduction to Numerical Analysis). We'll follow this pretty closely in lectures, though we will do the sections in reverse order!
- Stewart (*Afternotes ...*), Lectures 1–5. A well-presented introduction, with lots of diagrams to give an intuitive introduction.
- The proof of the convergence of Newton's Method is based on the presentation in Thm 3.2 of Epperson.

(The Root Finding Problem)

Let f be a continuous real-valued function defined on the interval $[a, b]$. Find $\tau \in [a, b]$ such that $f(\tau) = 0$.

Here f is some specified function, and τ is the **solution** to $f(x) = 0$.

This leads to three natural questions:

(1) Why bother?

(2) How do we know there is a solution?

(3) How do we find it?

Solving $f(x) = 0$ has been a mathematical goal for centuries. For example, it is the origin of all algebra.

But, although it appears quite specific, it is usually the simplest formulation of more general problem.

Suppose we want to find where two functions are equal:

. . . .

Many optimization problems are recast as root-finding:

2. When does a solution exist?

There are many root-finding problems (what is, choices of f, a and b) for which there is no solution, or (almost as bad) multiple solutions.

Example:

2. When does a solution exist?

The following gives *sufficient* conditions for the existence of a solution:

Theorem 2.1

Let f be a real-valued function that is defined and continuous on a bounded closed interval $[a, b] \subset \mathbb{R}$. Suppose that $f(a)f(b) \leq 0$. Then there exists $\tau \in [a, b]$ such that $f(\tau) = 0$.

3. Finding the solution

So now we know there is a solution τ to $f(x) = 0$, but how to we actually solve it? **Usually we don't!** Instead we construct a sequence of estimates $\{x_0, x_1, x_2, x_3, \dots\}$ that **converge** to the true solution. So now we have to answer these questions:

(1) How can we construct the sequence $x_0, x_1, \dots$?

(2) How do we show that $\lim_{k\to\infty} x_k = \tau$?

3. Finding the solution

There are some subtleties here, particularly with part (2). What we would like to say is that at each step the error is getting smaller. That is

$$|\tau - x_k| < |\tau - x_{k-1}| \qquad \text{for } k = 1, 2, 3, \ldots.$$

But we can't. Usually all we can say is that the **bounds** on the error is getting smaller. That is: **let ε_k be a bound on the error at step k**

$$|\tau - x_k| < \varepsilon_k,$$

then $\varepsilon_{k+1} < \mu\varepsilon_k$ for some number $\mu \in (0, 1)$. It is easiest to explain this in terms of an example, so we'll study the simplest method: **Bisection**.

Exercise 1.4.1. Does Theorem 1.1 mean that, if there is a solution to $f(x) = 0$ in $[a, b]$ then $f(a)f(b) \leq 0$? That is, is $f(a)f(b) \leq 0$ a *necessary* condition for their being a solution to $f(x) = 0$? Give an example that supports your answer.