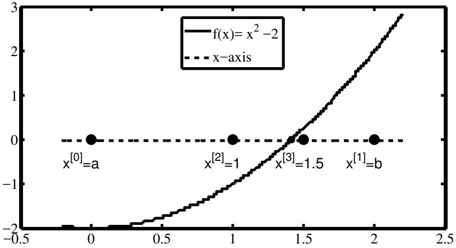
Start 16/9/2019.

Solving nonlinear equations

§1.1: The bisection method

MA385/530 – Numerical Analysis

September 2019



Linear equations are of the form:

$$\textit{find } x \textit{ such that } ax + b = 0 \quad \text{so} \quad x = \frac{-b}{a}.$$

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.
- Chapter 4 of Moler's "Numerical Computing with MATLAB". Gives a brief introduction to the methods we study, and description of MATLAB functions for solving these problems.
- The proof of the convergence of Newton's Method is based on the presentation in Thm 3.2 of Epperson.

Our generic problem is:

Let f be a continuous function 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 two natural questions:

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

(2) How do we find it?

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

(finished here w2.1)

Theorem 1.1

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 $f(\tau) = 0$.

Proof. If $\tau = a$ or $\tau = b$, then there is a solution. Otherwise $f(a)f(b) < 0$.

So $f(a)$ & $f(b)$ have different signs.

So, by the intermediate value theorem, there is some $\tau \in [a, b]$ with $f(\tau) = 0$.

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 $\lim_{k\to\infty} x_k = \tau$?

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***.

Bisection

The most elementary algorithm is the "*Bisection Method*" (also known as "Interval Bisection"). Suppose that we know that f changes sign on the interval $[a, b] = [x_0, x_1]$ and, thus, $f(x) = 0$ has a solution, τ, in $[a, b]$. Proceed as follows

1. Set x_2 to be the midpoint of the interval $[x_0, x_1]$.
2. Choose one of the sub-intervals $[x_0, x_2]$ and $[x_2, x_1]$ where f change sign;
3. Repeat Steps 1–2 on that sub-interval, until f is sufficiently small at the end points of the interval.

This may be expressed more precisely using some *pseudocode*.

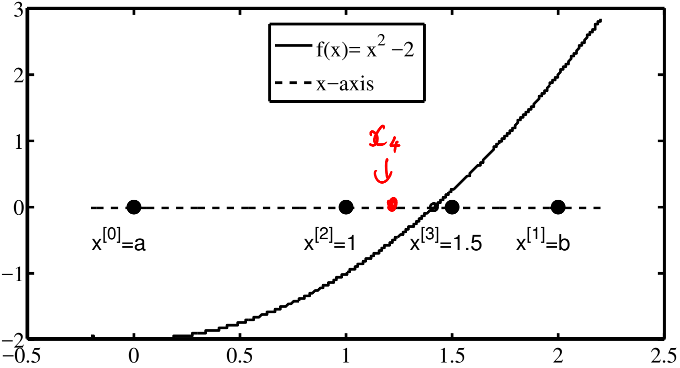
The Bisection Algorithm

Set eps to be the stopping criterion.
If $|f(a)| \leq eps$, return a. Exit.
If $|f(b)| \leq eps$, return b. Exit.
Set $x_L = a$ and $x_R = b$.
Set $k = 1$
while($|f(x_k)| > eps$)
    $x_{k+1} = (x_L + x_R)/2$;
    if $(f(x_L)f(x_{k+1}) < eps)$
        $x_R = x_{k+1}$;
    else
        $x_L = x_{k+1}$
    end if;
    $k = k + 1$
end while;

 $f(0)f(2) = (-2)(2) = -4.$

Example 1

Find an estimate for $\sqrt{2}$ that is correct to 6 decimal places.
Solution: Use bisection to solve $f(x) := x^2 - 2 = 0$ on the interval $[0, 2]$.



Example 1

Find an estimate for $\sqrt{2}$ that is correct to 6 decimal places.

Solution: Use bisection to solve $f(x) := x^2 - 2 = 0$ on the interval $[0, 2]$.

k	x_k	$\lvert x_k - \tau\rvert$	$\lvert x_k - x_{k-1}\rvert$
0	0.000000	1.41	
1	2.000000	5.86e-01	
2	1.000000	4.14e-01	1.00
3	1.500000	8.58e-02	5.00e-01
4	1.250000	1.64e-01	2.50e-01
5	1.375000	3.92e-02	1.25e-01
6	1.437500	2.33e-02	6.25e-02
7	1.406250	7.96e-03	3.12e-02
8	1.421875	7.66e-03	1.56e-02
9	1.414062	1.51e-04	7.81e-03
10	1.417969	3.76e-03	3.91e-03
⋮	⋮	⋮	⋮
22	1.414214	5.72e-07	9.54e-07

The main advantages of the Bisection method are

- It will always work.
- After k steps we know that

Theorem 1.2

$$|\tau - x_k| \le \left(\frac{1}{2}\right)^{k-1}|b-a|, \quad \text{for } k = 2, 3, 4, \ldots$$

Proof: x_1 is the midpoint of a and b. So $|\tau - x_2| \le \frac{1}{2}|b-a|$. Suppose $|\tau - x_k| \le \left(\frac{1}{2}\right)^{k-1}|b-a|$.

Then, since x_{k+1} bisects the interval to the left or right of x_k,

$$|\tau - x_k| \le \frac{1}{2}\left[\left(\frac{1}{2}\right)^{k-1}|b-a|\right].$$

So, by induction, ...

The main advantages of the Bisection method are

- It will always work.
- After k steps we know that

Theorem 1.2

$$|\tau - x_k| \le \left(\frac{1}{2}\right)^{k-1} |b-a|, \quad \text{for } k = 2, 3, 4, \ldots$$

note that, as $k \to \infty$, $(1/2)^{k-1} \to 0$.

So, $|\tau - x_k| \to 0$ as $k \to \infty$

so $x_k \to \tau$.

Improving upon bisection

A disadvantage of bisection is that it is not particularly efficient. So our next goal will be to derive better methods, particularly the ***Secant Method*** and ***Newton's method***. We also have to come up with some way of expressing what we mean by "better"; and we'll have to use Taylor's theorem in our analyses.

Exercise 1.1

Does Proposition 1.1.1 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.

Exercise 1.2

Suppose we want to find $\tau \in [a, b]$ such that $f(\tau) = 0$ for some given f, a and b. Write down an estimate for the number of iterations K required by the bisection method to ensure that, for a given ε, we know $|x_k - \tau| \leq \varepsilon$ for all $k \geq K$. In particular, how does this estimate depend on f, a and b?

Exercise 1.3

How many (decimal) digits of accuracy are gained at each step of the bisection method? (If you prefer, how many steps are needed to gain a single (decimal) digit of accuracy?)

Exercise 1.4

Let $f(x) = e^x - 2x - 2$. Show that there is a solution to the problem: *find* $\tau \in [0, 2]$ *such that* $f(\tau) = 0$.
Taking $x_0 = 0$ and $x_1 = 2$, use 6 steps of the bisection method to estimate τ. You may use a computer program to do this, but please note that in your solution.
Give an upper bound for the error $|\tau - x_6|$.

Exercise 1.5

We wish to estimate $\tau = \sqrt[3]{4}$ numerically by solving $f(x) = 0$ in $[a, b]$ for some suitably chosen f, a and b.

(i) Suggest suitable choices of f, a, and b for this problem.

(ii) Show that f has a zero in $[a, b]$.

(iii) Use 6 steps of the bisection method to estimate $\sqrt[3]{4}$. You may use a computer program to do this, but please note that in your solution.

(iv) Use Theorem 1.2 to give an upper bound for the error $|\tau - x_6|$.

Exercise 1.5

We wish to estimate $\tau = \sqrt[3]{4}$ numerically by solving $f(x) = 0$ in $[a, b]$ for some suitably chosen f, a and b.

(i) Suggest suitable choices of f, a, and b for this problem.

(ii) Show that f has a zero in $[a, b]$.

(iii) Use 6 steps of the bisection method to estimate $\sqrt[3]{4}$. You may use a computer program to do this, but please note that in your solution.

(iv) Use Theorem 1.2 to give an upper bound for the error $|\tau - x_6|$.