Nonlinear Equations (Chapter 5)

Problem: find x such that $f(x) = y$

where f is any smooth function.

(for now, $f: \mathbb{R} \to \mathbb{R}$)

Equivalently: find the root(s) of a general equation of the form $f(x) = 0$

subtract y from both sides

i.e. find an x^* such that $f(x^*) = 0$.

Existence and Uniqueness

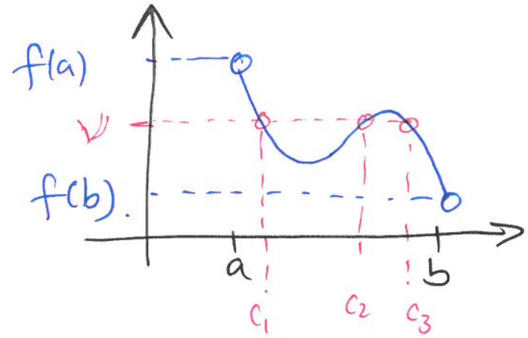
unlike in a linear system $f(x) = 0$ can have.

- no solution e.g. $e^x + 1 = 0$
- one solution e.g. $e^x - x = 0$.
- many solutions e.g. $x^3 + 6x^2 + 11x - 6 = 0$.
- inf. many solutions $\sin(x) - 1 = 0$

Even though we can't make any global assertions about existence / uniqueness of roots, there are useful local criteria that guarantees the existence of solutions.

Intermediate Value Thm

If f is continuous on an interval $[a, b]$, and ν is a number between $f(a)$ and $f(b)$, then there exists $c \in [a, b]$ such that $f(c) = \nu$.



$f(c_i) = v$.

We can apply I.V.T. to root finding:

> If f is continuous, $f(L)$ and $f(R)$ have opposite signs, then there is a root x^* between L and R with $f(x^*) = 0$.

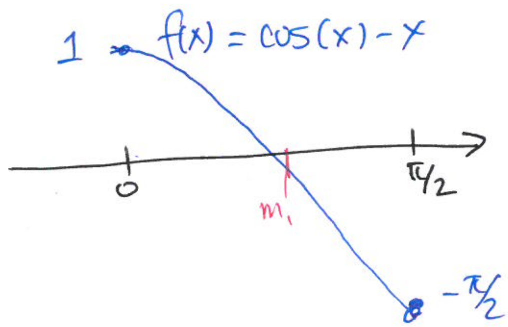
We can use this idea to develop an algorithm for root finding:

eg //. $f(x) = \cos(x) - x$.

$f(0) = 1 \qquad f(\frac{\pi}{2}) = -\frac{\pi}{2}$.

$\Rightarrow$ There should be a root $f(x^*)$

$x^* \in [0, \frac{\pi}{2}]$



Consider the midpoint $m_1 = 0 + \dfrac{\frac{\pi}{2} - 0}{2} = \dfrac{\pi}{4}$

we evaluate $f(m_1) = -0.07829$. ← negative.

Reapply the I.V.T. $\Rightarrow$ There should be a root $f(x^*)$

$f(0) > 0$

$f(m_1) < 0$

$x^* \in [0, \underset{m_1}{\underset{\shortparallel}{\frac{\pi}{4}}}]$

Consider the next midpoint $m_2 = 0 + \frac{\frac{\pi}{4} - 0}{2} = \frac{\pi}{8}$.

$$f(m_2) = 0.53118 \leftarrow \text{positive}$$

Reapply the IVT $\Rightarrow$ There should be a root $f(x^*)$

$f(m_2) > 0$
$f(m_1) < 0$

$$x^* \in \Big[\underbrace{\tfrac{\pi}{8}}_{= m_2}, \underbrace{\tfrac{\pi}{4}}_{= m_1}\Big]$$

⋮

<u>Algorithm</u> Interval Bisection.

We start with f continuous, and values a, b such that $\text{sign}(f(a)) \neq \text{sign}(f(b))$.

Iteratively find $m = a + \frac{(b-a)}{2}$.

If $\text{sign}(f(a)) = \text{sign}(f(m))$.
    $a := m$ new interval : $[m, b]$
else :
    $b = m$.   new interval : $[a, m]$.

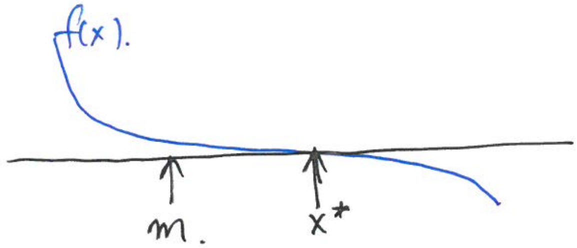
<u>Question</u>
* When do we stop?
* How quickly does Interval Bisection Converge?
* Can we do better? (Yes).
* What ~~can~~ we say about the <u>conditioning</u> of the root-finding problem $f(x) = 0$

Stopping Criteria

$|f(m)| < \varepsilon$

Idea #1: What about when $f(m)$ is small enough.

Can be fooled by flat functions:



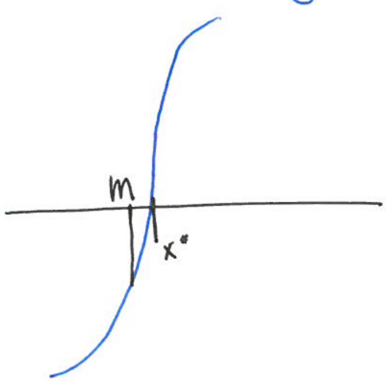
$f(m)$ is small

But $\underbrace{|m - x^*|}_{\text{absolute error}}$ large

Idea #2: What about when the interval size is sufficiently small? $\Rightarrow$ fixed number of iterations

$\sim |a_i - b_i| \le \varepsilon$

Can be fooled by functions like this:



$|x^* - m|$ is small

But $f(m)$ is large

In practise, the stopping criteria is (real life) problem dependent.

Convergence Rate

We use notation e_k to denote the absolute error

$e_k = x_k - x^*$ at iteration k.

where x_k is the approx. solution at iteration k.

x^* is the true solution.

In interval bisection ~~...~~ $|e_k| \le |b_k - a_k|$

we will approximate $e_k = |b_k - a_k|$.

An iterative method <u>converge with rate r.</u> if

$$\lim_{k\to\infty} \frac{\|e_{k+1}\|}{\|e_k\|^r} = C \qquad \text{for some constant } C>0.$$

In particular if

$r=1$, $C<1$. convergence rate is <u>linear</u>

$r>1$. " " " <u>superlinear</u>

$r=2$ <u>quadratic</u>

$r=3$ <u>cubic</u>

better (faster)

For interval bisection.

$|e_k| = |b_k - a_k|$ ← interval size at step k.

$= \dfrac{|b-a|}{2^k}$. ← ~~interval~~ initial interval size

So $\lim_{k\to 0} \dfrac{|e_{k+1}|}{|e_k|^1} =$ $\displaystyle\lim_{k\to\infty} \frac{|b-a|/2^{k+1}}{|b-a|/2^k} = \lim_{k\to\infty} \frac{1}{2} = \frac{1}{2}$

$\Rightarrow$ Interval Bisection converges <u>linearly</u>

with $r=1$, $c=\frac{1}{2}$.

$\Rightarrow$ we gain one additional correct bit of information in the approx solution per iteration

midpoint computation

Consider $F(\beta=10, p=2, U=10, L=-10)$.

$a = 0.67 \quad = \quad 6.7 \times 10^{-1}$

$b = 0.69 \quad = \quad 6.9 \times 10^{-1}$.

$m = \frac{a+b}{2}$ $\qquad a+b = 1.36 \Rightarrow 1.4 \times 10^{0}$.

$\frac{fl(a+b)}{2} = 0.70 \rightarrow 7.0 \times 10^{-1}$.

problem! $\quad m = 0.70$ is not even between $[a, b]$.

The computation $m = a + \frac{b-a}{2}$ guarantees that $m \in [a, b]$.

Fixed-Point Iteration

Find an equivalent problem to $f(x^*) = 0$ with the same solution.

↳ Find p^* that is a fixed-point of $g(x)$:
i.e. $g(p^*) = p^*$.

eg// Say we want to find a root of

$$f(x) = x - 0.2\sin(x) - 0.5$$

Define $g(x) = 0.2\sin(x) + 0.5$.

Notice that if $x^* = g(x^*)$ is a fixed-point of g.

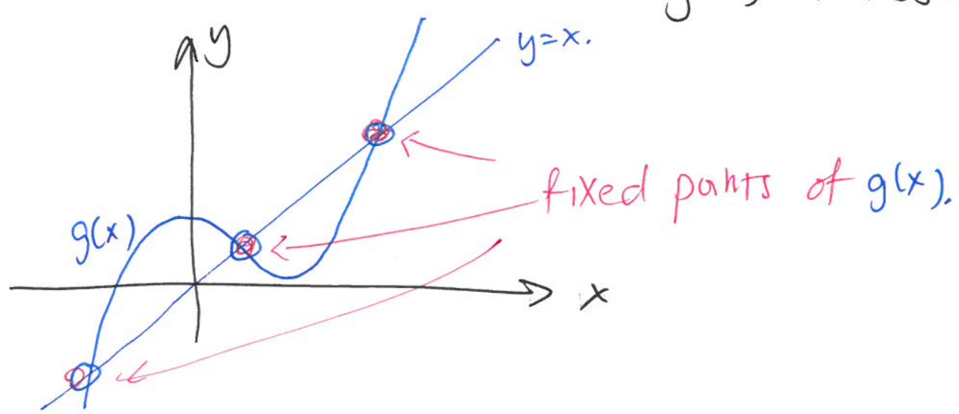
then $f(x^*) = 0$.

i.e. $x - \underbrace{0.2\sin(x) - 0.5} = 0$.

$\iff x = 0.2\sin(x) + 0.5$

Fixed-points of $g(x)$ are roots of $f(x)$.

Fixed-points are values where $g(x)$ intersects $y = x$.



We can find fixed-points of $g(x)$ using functional iteration (fixed-point iteration)

start with: x_0 given

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

If this iteration converges, it converges to a fixed point of g.

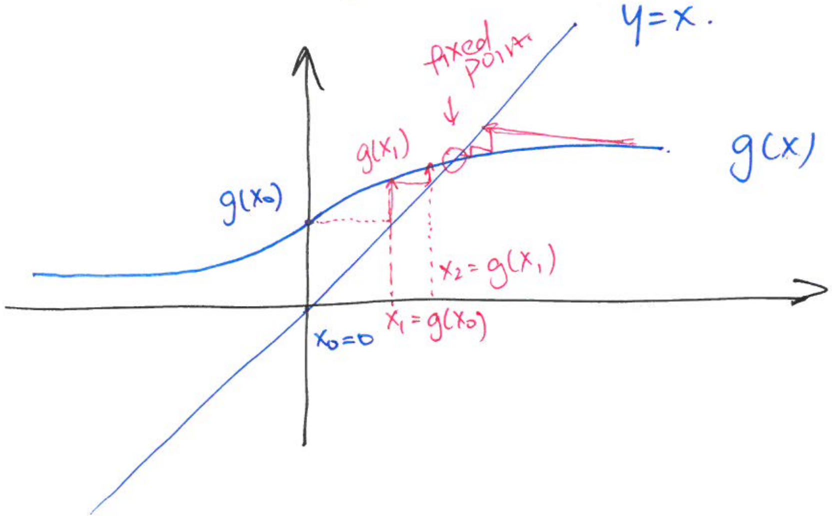
eg//. $f(x) = x - 0.2\sin(x) - 0.5$.

Run bisection $\Rightarrow$ ~22 iters to convergence

with $g(x) = 0.2\sin(x) + 0.5$.

start with $x_0 = 0$. $\Rightarrow$ ~8-9 iters.
$x_0 = 1.0$ $\Rightarrow$ ~8-9 iters
$x_0 = -1.0$ $\Rightarrow$ ~8-9 iters
$x_0 = 1000$
-1000 $\Rightarrow$ ~8-9 iters.

steps indep of x_0.

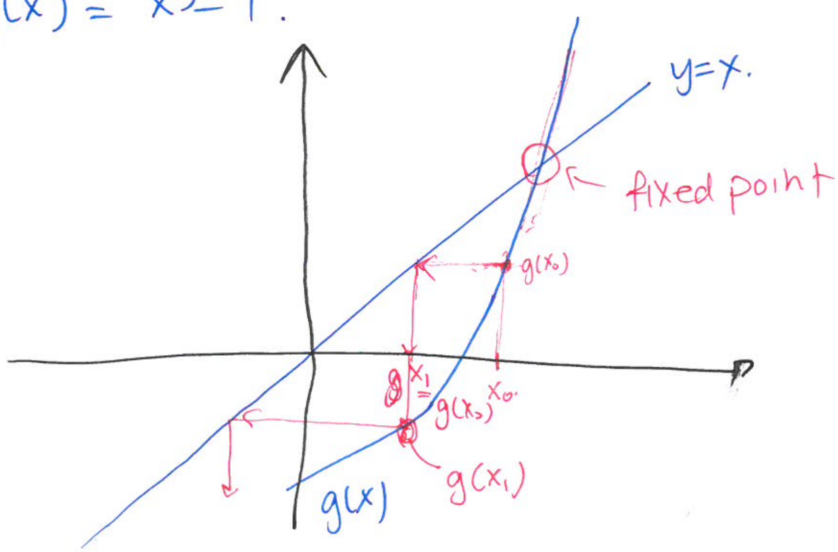


$g'(x) < 1$

eg// find roots of $f(x) = x^3 - x - 1$.

$x^3 - x - 1 = 0$
$x^3 - 1 = x$

$g(x) = x^3 - 1$.



$g'(x) > 1$.

Try $g(x) = (x+1)^{1/3}$.

$\Rightarrow$ Works!

$x^3 - x - 1 = 0$

$x = \sqrt[3]{x+1}$

Convergence of fixed-point iteration

Suppose $g(x)$ has a fixed point $x^* = g(x^*)$

If $|g'(x^*)| < 1$ then there is an interval containing x^* such that fixed-point iteration

$x_{k+1} = g(x_k)$ converges to x^* ~~for~~

if x_0 is in that interval

If $|g'(x^*)| > 1$. then iterative scheme diverges.

The asymptotic convergence rate is usually linear with $C = |g'(x^*)|$

If $g'(x^*) = 0$ then the convergence is at least quadratic