Lecture 6

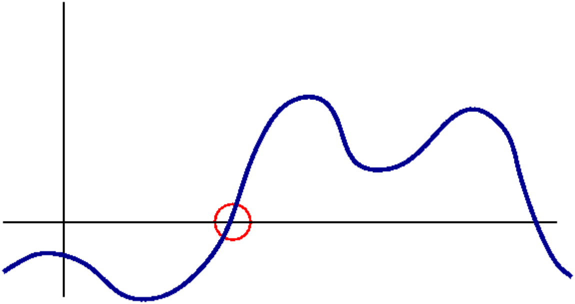
Nonlinear Equations: Newton and Secant Methods

Owen L. Lewis

Department of Mathematics and Statistics
University of New Mexico

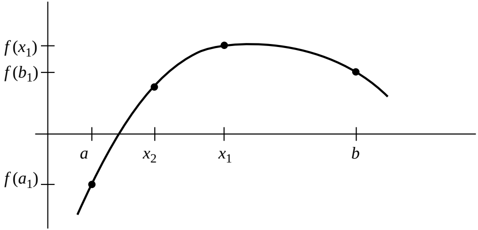
Last Time: Root Finding

Given a function $f(x)$, find x so that $f(x) = 0$



Bisection

Given a bracketed root, halve the interval while continuing to bracket the root



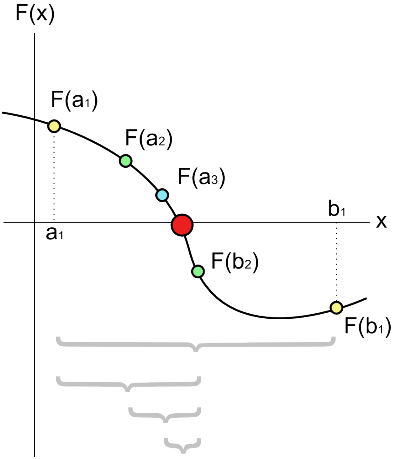
Bisection (2)

For the bracket interval $[a, b]$ the midpoint is

$$x_m = \frac{1}{2}(a + b)$$

Idea:

1. split bracket in half
2. select the bracket that has the root
3. goto step 1



Analysis of Bisection

$$\frac{\delta_n}{\delta_0} = \left(\frac{1}{2}\right)^n = 2^{-n} \qquad \text{or} \qquad n = -\log_2\left(\frac{\delta_n}{\delta_0}\right)$$

The ratio $\frac{\delta_n}{\delta_0}$ measures a relative reduction of your error (assuming you guess that the root is somewhere inside the bracketed interval)

n	$\frac{\delta_n}{\delta_0}$	function evaluations
5	3.1×10^{-2}	7
10	9.8×10^{-4}	12
20	9.5×10^{-7}	22
30	9.3×10^{-10}	32
40	9.1×10^{-13}	42
50	8.9×10^{-16}	52

Bisection: Error in Root

- If we pick the midpoint $c_n = (a_n + b_n)/2$ as the root then,

$$|x_* - c_n| \leqslant (b_n - a_n)/2 = \delta_n/2$$

  where x_* is the true root.
- Recall that

$$\delta_n = \left(\frac{1}{2}\right)^n \delta_0$$

- The error in the root after n steps is

$$\begin{aligned} |x_* - c_n| \leqslant (b_n - a_n)/2 &= \left(\frac{1}{2}\right)^{n+1} \delta_0 \\ &= \left(\frac{1}{2}\right)^{n+1} (b - a) \end{aligned}$$

Bisection: Example

Question: How many steps of bisection are needed in order to compute the root of f so that the error is less than 10^{-8} if $a = -2$ and $b = 3$?

Solution: Find n such that,

$$
\begin{aligned}
& \left(\frac{1}{2}\right)^{n+1} (b-a) \leqslant 10^{-8} \\
\implies & \left(\frac{1}{2}\right)^{n+1} (5) \leqslant \times 10^{-8} && \text{multiply by } 2/5 \text{, and take natural log} \\
\implies & \log 2^{-n} \leqslant \log((2/5) \times 10^{-8}) && \text{divide through by} (-\log 2) \\
\implies & n \geqslant \frac{1}{\log 2}(-\log(2/5) + 8 \log 10) && \text{just evaluate expressions} \\
\implies & n \geqslant 27.89
\end{aligned}
$$

So, want $n \geqslant 28$ steps.

Is this “good enough”?

This will get us an approx x_n and it is within 10^{-8} of the true answer x^*.

Is it good enough?

Is $f(x_n)$ close to zero?

Convergence Criteria

An automatic root-finding procedure needs to monitor progress towards the root and then stop when the current guess is close enough to the desired root.

This will avoid unneccessary work. Two convergence checks are:

- Check the closeness of successive approximations, for tolerance δ_x

$$|x_n - x_{n-1}| < \delta_x$$

- Check how close $f(x)$ is to zero at the current guess, for tolerance δ_f

$$|f(x_n)| < \delta_f$$

- Which one you use depends on the problem being solved

Different Criteria

- How close we are to the “true” answer:
  Tolerance on input, forward error (book), “error”.

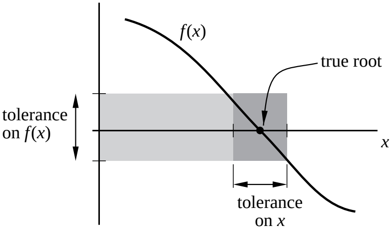
$$|x_n - x^*|$$

- How close we are to solving the problem:
  Tolerance on output, backward error (book), “residual”

$$|f(x_n) - 0|$$

- They are related, but *NOT* the same.

Convergence Criteria on x

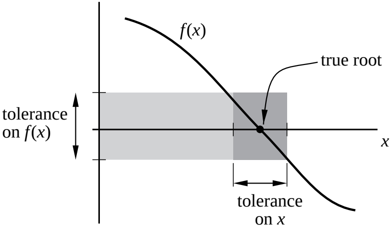


$$x_n = \text{current guess at the root (midpoint of current bracket)}$$
$$x_{n-1} = \text{previous guess at the root (midpoint of previous bracket)}$$

Absolute tolerance: $\left|x_n - x_{n-1}\right| < \delta_x$

Relative tolerance: $\left|\dfrac{x_n - x_{n-1}}{b - a}\right| < \hat{\delta}_x$

Convergence Criteria on $f(x)$



Absolute tolerance: $\left|f(x_n)\right| < \delta_f$

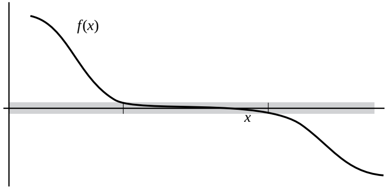
Relative tolerance:

$$|f(x_n)| < \frac{\hat{\delta}_f}{\max\left\{|f(a_0)|, |f(b_0)|\right\}}$$

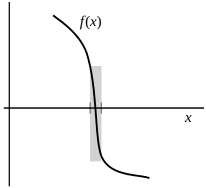
where a_0 and b_0 are the original brackets

Convergence Criteria Compared

If $f'(x)$ is small near the root, it is easy to satisfy tolerance on $f(x)$ for a large range of Δx. The tolerance on Δx is more conservative (safer)



If $f'(x)$ is large near the root, it is possible to satisfy the tolerance on Δx when $|f(x)|$ is still large. The tolerance on $f(x)$ is more conservative (safer)



Relationship Between Criteria

- How are the criteria on x and $f(x)$ related? Consider the ratio of the two criteria
$$\frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}$$
- The limit of this as x_{n-1} and x_n converge to the exact answer x^* is just $f'(x^*)$.
- We can thus expect (this is not yet a proof) that
$$|f(x_n) - f(x_{n-1})| \approx |f'(x^*)||x_n - x_{n-1}|$$
as x_{n-1} and x_n approach the solution x^*.

Convergence rate of a root finding iteration

- Let $e_n = x^* - x_n$ be the error.
- In general, a sequence is said to converge with rate r if

$$\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|^r} \leqslant C$$

Special Cases:

- If $r = 1$ and $C < 1$, then the rate is *linear*
- If $r = 2$ and $C > 0$, then the rate is *quadratic*
- If $r = 3$ and $C > 0$, then the rate is *cubic*

Example

Convergence Rate

1. $10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}...$
2. $10^{-2}, 10^{-4}, 10^{-6}, 10^{-8}...$
3. $10^{-2}, 10^{-4}, 10^{-8}, 10^{-16}...$
4. $10^{-2}, 10^{-6}, 10^{-18}, ...$

Example

Convergence Rate

1. $10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}$... (linear with $C = 10^{-1}$)
2. $10^{-2}, 10^{-4}, 10^{-6}, 10^{-8}$... (linear with $C = 10^{-2}$)
3. $10^{-2}, 10^{-4}, 10^{-8}, 10^{-16}$...(quadratic)
4. $10^{-2}, 10^{-6}, 10^{-18}$, ... (cubic)

- Linear: Adds a constant amount of accuracy at each step (say one digit, or two digits of accuracy)
- Quadratic: Adds double the number of digits of accuracy (relative to the previous step) at each step

Convergence rate of Bisection Algorithm

The worst case of Bisection algorithm is,

$$|e_n| = |x_* - c_n| = \left(\frac{1}{2}\right)^{n+1} \delta_0$$

And so, in the worst case we have,

$$\frac{|e_{n+1}|}{|e_n|} = \frac{1}{2}$$

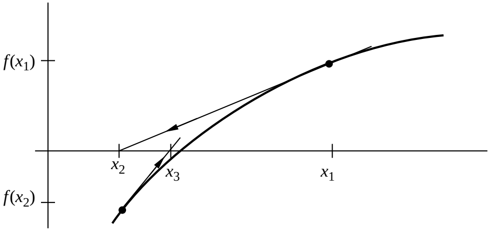
That is, bisection converges linearly with $C = 1/2$.

Convergence rate of Bisection Algorithm

Bisection converges linearly with $C = 1/2$.

Can we do better? Newton and Secant methods.

Newton's Method



For a current guess x_k, use $f(x_k)$ and the slope $f'(x_k)$ to predict where $f(x)$ crosses the x axis.

Notes

Newton's Method: Derivation

Expand Taylor Series of $f(x)$ around x_k

$$f(x_k + \Delta x) = f(x_k) + \Delta x \left.\frac{df}{dx}\right|_{x_k} + \frac{(\Delta x)^2}{2} \left.\frac{d^2 f}{dx^2}\right|_{x_k} + \ldots$$

Substitute $\Delta x = x_{k+1} - x_k$
and neglect 2^{nd} order terms to get

$$f(x_{k+1}) \approx f(x_k) + (x_{k+1} - x_k)\, f'(x_k)$$

where

$$f'(x_k) = \left.\frac{df}{dx}\right|_{x_k}$$

This is a *linear* approximation to $f(x)$, for x close to x_k. Find the zeros of this approximation!

Newton's Method: Derivation

Goal is to find x such that $f(x) = 0$.

Set $f(x_{k+1}) = 0$ and solve for x_{k+1}. Remember,

$$f(x_{k+1}) \approx f(x_k) + (x_{k+1} - x_k)\, f'(x_k)$$

Then,

$$0 = f(x_k) + (x_{k+1} - x_k)\, f'(x_k)$$

or, solving for x_{k+1}

$$\boxed{x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}}$$

It's really that simple...

Newton's Method: Algorithm

```
initialize:  x_1 = ...
for k = 2, 3, ...
   x_k = x_{k-1} - f(x_{k-1})/f'(x_{k-1})
   if converged, stop
end
```

Newton's Method: Example

Solve:

$$x - x^{1/3} - 2 = 0$$

First derivative is

$$f'(x) = 1 - \frac{1}{3}x^{-2/3}$$

The iteration formula is

$$x_{k+1} = x_k - \frac{x_k - x_k^{1/3} - 2}{1 - \frac{1}{3}x_k^{-2/3}}$$

Newton's Method: Example

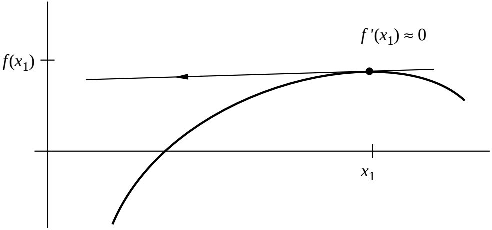
$$x_{k+1} = x_k - \frac{x_k - x_k^{1/3} - 2}{1 - \frac{1}{3}x_k^{-2/3}}$$

k	x_k	$f'(x_k)$	$f(x)$
0	3	0.83975005	-0.44224957
1	3.52664429	0.85612976	0.00450679
2	3.52138015	0.85598641	3.771×10^{-7}
3	3.52137971	0.85598640	2.664×10^{-15}
4	3.52137971	0.85598640	0.0

Conclusion

- Newton's method converges *much* more quickly than bisection
- Newton's method requires an analytical formula for $f'(x)$
- The algorithm is simple, as long as $f'(x)$ is available.
- Iterations are not guaranteed to stay inside an initial bracket.

Divergence of Newton's Method



Since

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

the new guess, x_{k+1}, will be far from the old guess whenever $f'(x_k) \approx 0$

Newton's Method: Convergence

Recall

Convergence of a method is said to be of order r if there is a constant $C > 0$

$$\lim_{k\to\infty} \frac{|e_{k+1}|}{|e_k|^r} \leqslant C$$

Newton's method is of order 2 (quadratic) when $f'(x_*) \neq 0$.

Proof: For some ξ between x_k and x_*,

$$f(x_*) = f(x_k) + (x_* - x_k)f'(x_k) + \frac{1}{2}(x_* - x_k)^2 f''(\xi) = 0$$

So

$$\frac{f(x_k)}{f'(x_k)} + x_* - x_k + (x_* - x_k)^2 \left(\frac{1}{2}\right) \frac{f''(\xi)}{f'(x_k)} = 0$$

Then

$$\left(\frac{f(x_k)}{f'(x_k)} - x_k\right) + x_* + (x_* - x_k)^2 \left(\frac{1}{2}\right) \frac{f''(\xi)}{f'(x_k)} = 0$$

$$x_* - x_{k+1} + (x_* - x_k)^2 \left(\frac{1}{2}\right) \frac{f''(\xi)}{f'(x_k)} = 0$$

Thus

$$\frac{|x_* - x_{k+1}|}{|x_* - x_k|^2} = \left(\frac{1}{2}\right) \left|\frac{f''(\xi)}{f'(x_k)}\right|$$

Newton's Method: Convergence – Single and Multiple Roots

Proof: Cont'd

- Thus, for $f'(x_*) \neq 0$ we have,

$$\frac{|e_{k+1}|}{|e_k|^2} = \frac{|x_* - x_{k+1}|}{|x_* - x_k|^2} = \left(\frac{1}{2}\right)\left|\frac{f''(\xi)}{f'(x_k)}\right|$$

  Near the root x_* we can bound $|\frac{f''(\xi)}{f'(x_k)}| < C$. (See Pg 130 in NMC7 for details).
  In this case, Newton iteration converges **quadratically**.

- What if $f'(x_*) = 0$?
  Multiple root or multiple zero of f.

  In this case, Newton iteration converges only **linearly**!

Secant Method

Given

$$\begin{aligned} x_k &= \text{current guess at the root} \\ x_{k-1} &= \text{previous guess at the root} \end{aligned}$$

Approximate the first derivative with

$$f'(x_k) \approx \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}$$

Substitute approximate $f'(x_k)$ into formula for Newton's method

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

to get

$$x_{k+1} = x_k - f(x_k)\left[\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}\right]$$

Secant Method

Two versions of this formula are (equivalent in exact math)

$$x_{k+1} = x_k - f(x_k)\left[\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}\right] \qquad (\star)$$

and

$$x_{k+1} = \frac{f(x_k)x_{k-1} - f(x_{k-1})x_k}{f(x_k) - f(x_{k-1})} \qquad (\star\star)$$

Equation $(\star)$ is better since it is of the form $x_{k+1} = x_k + \Delta$. Even if Δ is inaccurate the change in the estimate of the root will be small at convergence because $f(x_k)$ will also be small.

Equation $(\star\star)$ is susceptible to catastrophic cancellation:

- $f(x_k) \to f(x_{k-1})$ as convergence approaches, so cancellation error in denominator can be large.
- $|f(x)| \to 0$ as convergence approaches, so underflow is possible in the numerator of $(\star\star)$

Secant Algorithm

```
initialize:  $x_1 = \ldots$, $x_2 = \ldots$
for $k = 2, 3 \ldots$
  $x_{k+1} = x_k - f(x_k)(x_k - x_{k-1})/(f(x_k) - f(x_{k-1}))$
  if converged, stop
end
```

Newton's Method: Example

Solve:

$$x - x^{1/3} - 2 = 0$$

Beginning with the interval $[3, 4]$

k	x_k	$f'(x_k)$	$f(x)$
0	3	0.83975005	-0.44224957
1	3.52664429	0.85612976	0.00450679
2	3.52138015	0.85598641	3.771×10^{-7}
3	3.52137971	0.85598640	2.664×10^{-15}
4	3.52137971	0.85598640	0.0

Secant Example

Solve:

$$x - x^{1/3} - 2 = 0$$

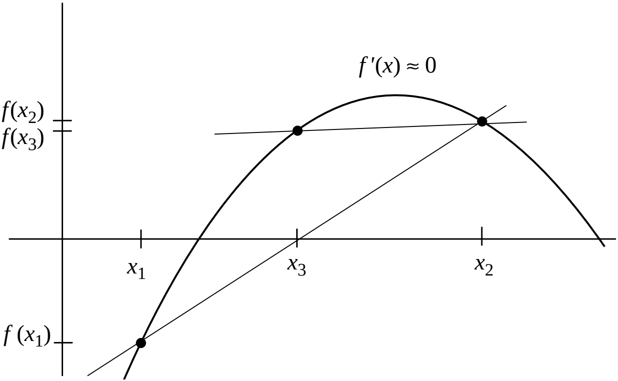
Beginning with the interval $[3, 4]$

k	x_{k-1}	x_k	$f(x_k)$
0	4	3	-0.44224957
1	3	3.51734262	-0.00345547
2	3.51734262	3.52141665	0.00003163
3	3.52141665	3.52137970	-2.034×10^{-9}
4	3.52137959	3.52137971	-1.332×10^{-15}
5	3.52137971	3.52137971	0.0

Conclusions:

- Converges almost as quickly as Newton's method ($r \approx 1.62$).
- There is no need to compute $f'(x)$.
- The algorithm is simple.
- Two initial guesses are necessary
- Iterations are not guaranteed to stay inside an original bracket.

Divergence of Secant Method



Since

$$x_{k+1} = x_k - f(x_k)\left[\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}\right]$$

the new guess, x_{k+1}, will be far from the old guess whenever $f(x_k) \approx f(x_{k-1})$ and $|f(x)|$ is not small.

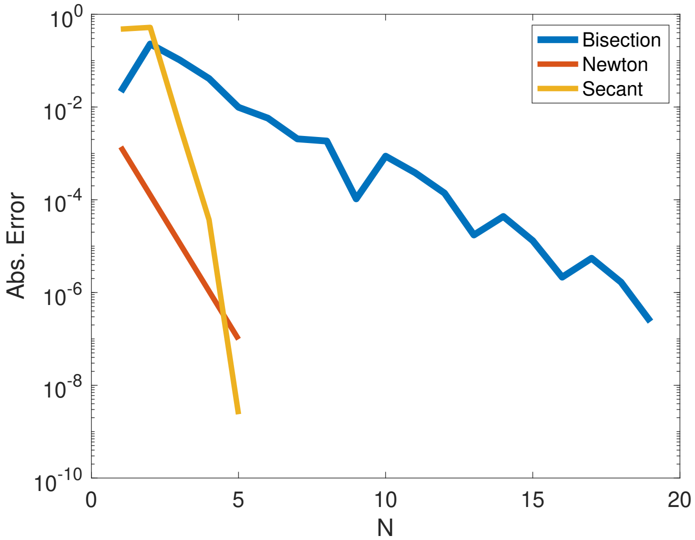
This issue is particularly bad if $f'(x_*) \approx 0$

Summary

- Plot $f(x)$ before searching for roots
- Bracketing finds coarse interval containing roots and singularities
- Bisection is robust, but converges slowly
- Newton's Method
  - Requires $f(x)$ and $f'(x)$.
  - Iterates are not confined to initial bracket.
  - Converges rapidly ($r = 2$).
  - Diverges if $f'(x) \approx 0$ is encountered.
- Secant Method
  - Uses $f(x)$ values to approximate $f'(x)$.
  - Iterates are not confined to initial bracket.
  - Converges almost as rapidly as Newton's method ($r \approx 1.62$).
  - Diverges if $f'(x) \approx 0$ is encountered.

Notes

Error Plots



Bisection
Newton
Secant
10^{0}
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
Abs. Error
0
5
10
15
20
N