Root Finding

Goal: Solve $f(r) = 0$. The solution r is called a *root.*

Do we need methods for solving $f(x) = c$ for $c \neq 0$?

Remark. No. Subtract c from both sides and you have a new function, $f(x) - c$, whose roots are the solutions to $f(x) = c$.

You've had the most experience solving polynomials like $x^2 - 5x - 6 = 0$. Name some strategies! What would work best on floating point numbers?

Remark. We emphasize factoring in algebra, but there's a lot of quadratics that can't be easily factored, a fact which only gets worse when coefficients are now floating point numbers. The quadratic formula, however, does work on a computer... though the fact that we're discussing numerical problems with it in your homework isn't a good sign for its usefulness. There are also formulas for cubics and quartics, although the fact you've never been taught them should tell you something about how not-nice they are to evaluate. Using a formula for the roots is an example of a *direct* method - performing a finite number of operations produces the answer. But, it's been proven that no such formulas can exist for polynomials of degree 5 and higher…and polynomials are the NICE functions…

Since there's no way to immediately calculate the roots of an arbitrary function, we turn to *iterative* methods.

Definition 1. An iterative method consists of the following general process.

1. Start with a guess for the solution
2. Do something to get a new guess (hopefully better)
3. Repeat until a *stopping criteria* is met

If repeating this process gets you close to a correct answer, the method is called *convergent.*

Many common math problems can famously **only** be solved by iterative methods. General root finding is one of them. The eigenvalue problem from linear algebra is also one, which is a consequence of the root-finding problem. Eigenvalues are the roots of the characteristic equation (a polynomial)!

We are covering 4 iterative methods for root-finding: Bisection, Fixed Point Iteration, Newton's Method, and Secant Method. Why more than one? In short: there are differences in convergence, speed and accuracy that depend on the problem input.

Bisection Method

How do you even know if a function has a root?

Theorem 2 (Intermediate Value Theorem)**.** *Let f be a continuous function on $[a, b]$ satisfying $f(a)f(b) < 0$. Then f has a root between a and b; that is, there exists a number r satisfying $a < r < b$ and $f(r) = 0$.*

The intermediate value theorem is an example of a *sufficient* condition for having a root, but not a *necessary* condition.

Such an interval $[a, b]$ is said to *bracket* the root. The "best" guess for the root, given such an interval, is the midpoint $\frac{a+b}{2}$, where "best" means the smallest bound on the worst error. What is this maximum error for the midpoint?

Remark. The distance from the midpoint to the ends of the interval is $\frac{b-a}{2}$. This is the bound on the absolute error when using the midpoint as an approximation to the root.

Algorithm 1 Bisection Method Algorithm

```
Input: Function $f$, initial interval $[a, b]$, desired accuracy TOL
Check: $f(a)f(b) < 0$
while $\frac{b-a}{2} > TOL$ do
    New guess: $c = \frac{b+a}{2}$
    Compute $f(c)$
    if $f(a)f(c) < 0$ then
        $b = c$
    else if $f(a)f(c) > 0$ then
        $a = c$
    else
        $c$ is a root! Stop immediately
Return $\frac{b+a}{2}$
```

Example 3. Let $f(x) = 2x^2 - 1$. Apply three steps of the bisection method on $[0, 4]$.

Remark. First, note that $f(0)f(4) = (-1)(31) = -31 < 0$. Since f is a polynomial, it is continuous, and there IS a root in the given interval.

Then the initial guess is $r_0 = \frac{0+4}{2} = 2$ where $f(2) = 7$. Since $f(0)f(2) = -7 < 0$, we assign the new interval to be $[0, 2]$ and the new guess is $r_1 = \frac{0+2}{2} = 1$.

Since $f(0)f(1) = -1 < 0$, we assign the new interval to be $[0, 1]$ and the new guess is $r_2 = \frac{0+1}{2} = \frac{1}{2}$.

For the final step, note that $f(0)f(\frac{1}{2}) = (-1)(-\frac{1}{2}) > 0$. So the new interval is $[\frac{1}{2}, 1]$ and the final estimate for the root is $r_3 = \frac{3}{4}$.

Example 4. If you perform n steps of the bisection method on an interval $[a, b]$, how many times do you evaluate the function? What is the maximum error of the result?

Remark. For n iterations, maximum error is $\frac{b-a}{2^{n+1}}$. Performing n steps requires $n+2$ function evaluations.

Definition 5. A solution is *correct within p decimal places* if the error is less than 0.5×10^{-p}.

Is the result in Example 3 correct within 1 decimal place? If not, how many additional iterations would be needed?

Remark. We only know that the error is at most $\frac{4}{2^{3+1}} = \frac{1}{4}$. Since $\frac{1}{4} > \frac{1}{20}$, it is not correct to one decimal place. In three more iterations the maximum error is $\frac{1}{32} < \frac{1}{20}$. That means a total of six iterations is needed to go from knowing there's a root in $[0, 4]$ to starting to generate correct digits of the root. Five isn't awesome, but definitely not terrible either considering the initial interval was pretty big. We try to start with a unit interval (meaning $b = a + 1$) which shortens this initial number of iterations slightly.

Definition 6. Let e_i denote the (forward) error at step i of an iterative method ($e_i = |r - x_i|$). If

$$\lim_{i\to\infty} \frac{e_{i+1}}{e_i} = S < 1,$$

the method is said to obey *linear convergence* with rate S.

Does the bisection method converge linearly?

Remark. Yes, it does. The maximum error at step i is $\frac{b-a}{2^{i+1}}$, so we have $\lim_{i\to\infty} \frac{\frac{b-a}{2^{i+2}}}{\frac{b-a}{2^{i+1}}} = \lim_{i\to\infty} (\frac{1}{2^{i+1}(2)})(\frac{2^{i+1}}{1}) = \frac{1}{2}$. So the bisection method converges linearly with rate $\frac{1}{2}$. In the grand scheme of things, linear convergence isn't that impressive on speed, but it could be worse!

Fixed-Point Iteration

Definition 7. A real number x is a *fixed point* of the function g if $g(x) = x$.

Can fixed points be found by the Bisection Method?

Remark. Yes, certainly, by applying it to the function $g(x) - x$. We can go to the other way too: to find a root $g(x) = 0$, you can find fixed points of $g(x) + x$. So these problems are equivalent.

Algorithm 2 Fixed-point Iteration Algorithm (FPI)

```
Input: A function $f$, an initial guess $x_0$, number of iterations $n$
for $i = 1, \ldots, n$ do
    $x_i = g(x_{i-1})$
Return $x_n$
```

Example 8. Apply two steps of FPI to $f(x) = x^2$ with initial guess $\frac{1}{2}$. What fixed point are you approximating?

Remark.

$$\begin{aligned} x_0 &= \frac{1}{2} \\ x_1 &= f(\frac{1}{2}) = \frac{1}{4} \\ x_2 &= f(\frac{1}{4}) = \frac{1}{16} \end{aligned}$$

This is converging to the fixed point 0 (which you can find by solving $x^2 = x$ in the usual factoring manner). Note that the initial guess was exactly halfway between the two fixed points of the function, 0 and 1. Would any initial guess, other than 1 itself, converge to the fixed point at 1?

Draw a *cobweb diagram* showing the FPI process for this example.

Remark. See the posted video for this answer.

Definition 9. An iterative method is called *locally convergent* to a solution if the method converges for initial guesses sufficiently close to the solution. If it's convergent for any initial guess, the method is *globally convergent*.

FPI for Example 8 is locally convergent for the fixed point 0 as long as the initial guess is in $(-1, 1)$. It is not locally convergent for the fixed point 1, and it is not globally convergent.

Theorem 10. *Let $g(x)$ be a continuously differentiable function, r be a fixed point of g, and suppose that $|g'(r)| = S < 1$. Then Fixed-Point Iteration converges linearly with rate S to the fixed point r for initial guesses sufficiently close to r.*

Example 11. Use this theorem to predict the previously observed convergence regarding fixed points 0 and 1.

Remark. The derivative for this function is $f'(x) = 2x$. At $x = 0$, we have $2(0) = 0 < 1$, so this is locally convergent. But at $x = 1$, we have $2(1) = 2 > 1$, and thus the method is divergent here.

Example 12. What are the implications of Theorem 10 for sine and cosine?

Remark. Sine and Cosine's derivatives (being each other) are bounded in absolute value by 1. They should be locally convergent to their fixed points, though you should be aware that strange things can happen when the derivative equals one (like x and $x + 1$).

In FPI, we gave as input a pre-set number of iterations for the stopping criteria. But there is one alternative: to go until the answer doesn't seem to change much anymore. In other words, if $|x_i - x_{i-1}| < TOL$.

Remark. This can also be used as a failure detection point, meaning if $|x_i - x_{i+1}|$ is too large, then stop.