Homework 6 (root finding)

Hanzhang Yin

November 20, 2024

1. One view of the secant method: it is a coarser Newton's method. We've seen that it has some of the speed of Newton's method. One might also hope that it enjoys similar convergence properties.

Adapt the convergence proof for Newton's method to show that the secant method also always converges under the following assumptions about the function f on the interval $[a, b]$:

i) f is twice continuously differentiable

ii) $f' > 0$

iii) $f'' > 0$

iv) f has a root x in the interval

v) the two initial guesses x_0, x_1 are both to the right of the root.

Hint: you will have to use convexity in a slightly more interesting way than in NM – the graph of f does not lie above the secant line, but you can argue that the right (well, left!) piece still does.

Solution.

Proof. The secant method iterates according to the formula:

$$x_{n+1} = x_n - f(x_n) \cdot \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}.$$

We proceed in steps to show the convergence.

Monotonicity and Boundedness

Claim: The sequence $\{x_n\}$ is strictly decreasing and bounded below by x^*.

We will prove the claim by induction

- **Base Case ($n = 1$):** By assumption, $x_0 > x^*$ and $x_1 > x^*$. WLOG, we can reorder x_0 and x_1 such that $x_0 > x_1 > x^*$. Hence, the base case holds.
- **Inductive Step:** Assume $x_{n-2} > x_{n-1} > x^*$. We show that $x_n > x^*$ and $x_n < x_{n-1}$.

  From the secant update:

  $$x_n = x_{n-1} - f(x_{n-1}) \cdot \frac{x_{n-1} - x_{n-2}}{f(x_{n-1}) - f(x_{n-2})}.$$

  – Since $x_{n-2} > x_{n-1}$, we have $x_{n-1} - x_{n-2} < 0$.
  – Since $f'(x) > 0$ on $[a, b]$, $f(x_{n-1}) > f(x_{n-2})$, so $f(x_{n-1}) - f(x_{n-2}) > 0$.
  – Therefore, the ratio $\frac{x_{n-1} - x_{n-2}}{f(x_{n-1}) - f(x_{n-2})} < 0$.
  – Since $f(x_{n-1}) > 0$ (as $x_{n-1} > x^*$ and f is increasing), the term subtracted from x_{n-1} is positive:

  $$x_n = x_{n-1} - (\text{Positive Number}) < x_{n-1}$$

  implying $x_n < x_{n-1}$.
  – To show $x_n > x^*$, assume $x_n \leq x^*$. Then $f(x_n) \leq f(x^*) = 0$, contradicting the fact that $f(x_n) > 0$ for $x_n > x^*$. Thus, $x_n > x^*$.

By induction, $\{x_n\}$ is strictly decreasing and bounded below by x^*.
Convergence of the Sequence
Claim: The sequence $\{x_n\}$ converges to x^*.
Since $\{x_n\}$ is strictly decreasing and bounded below by x^*, it converges to some limit $l \geq x^*$ by **MCT**. Suppose, for contradiction, that $l > x^*$.

- Since f is continuous and strictly increasing:

$$\lim_{n \to \infty} f(x_n) = f(l) > f(x^*) = 0.$$

- Consider the secant update:

$$x_{n+1} = x_n - \frac{f(x_n)}{s_n}, \quad \text{where } s_n = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}.$$

- Because f is convex ($f'' > 0$), the slope $s_n > f'(x^*) > 0$, so:

$$\left| \frac{f(x_n)}{s_n} \right| < \frac{f(x_n)}{f'(x^*)}.$$

- As $n \to \infty$, $f(x_n) \to f(l) > 0$, meaning the step sizes $x_n - x_{n+1}$ do not shrink to zero.
- This contradicts the convergence $x_n \to l$, as the step sizes must tend to zero for convergence.

Thus, $l = x^*$, and the sequence converges to x^*.
The convexity of f ensures that the secant line between any two points lies below the graph of f, preventing the iterates x_n from overshooting the root x^*. Thus, $x_n > x^*$ for all n. Under the given assumptions:

1. $\{x_n\}$ is strictly decreasing and bounded below by x^*,
2. By the monotone convergence theorem, $\{x_n\}$ converges to a limit $l \geq x^*$,
3. Assuming $l > x^*$ leads to a contradiction, hence $l = x^*$,
4. Convexity ensures no overshooting, maintaining $x_n > x^*$.

Therefore, the secant method converges to the root x^*. $\square$

2. Another view of the secant method, discussed in class, is as a weighted bisection method. Here too, one might hope for a convergence guarantee, because BM is much more robust than NM in that regard.

Consider a modified secant method which at step k takes in endpoints a_k, b_k, calculates their weighted midpoint c_k and then returns two new endpoints a_{k+1}, b_{k+1}, one of which is c_k, to which IVT applies. These new endpoints are input to the next step.

Prove that if f is continuous on $[a, b] = [a_0, b_0]$ and the IVT applies to f on the interval, then the sequence c_k from the modified secant method converges to a root of f.

Hint: the reason for convergence is *not* the same as for bisection. This would require the stronger assumption that f is continuously differentiable. In fact:
[Bonus] Give an example where the sequences x_k and y_k converge to different points, so squeeze does not apply.

Solution.

Proof. From Bisection Method, at each iteration k, we have:

1. $f(a_k) \cdot f(b_k) < 0$, so f changes sign on $[a_k, b_k]$.
2. The point c_k is computed using the secant method formula:
$$c_k = b_k - f(b_k) \cdot \frac{b_k - a_k}{f(b_k) - f(a_k)}.$$
3. One of the endpoints a_{k+1} or b_{k+1} is set to c_k, and the other remains a_k or b_k, such that $f(a_{k+1}) \cdot f(b_{k+1}) < 0$.
4. *The intervals $[a_k, b_k]$ may not necessarily shrink, but they always contain a root of f due to the IVT.

Since $\{c_k\}$ is bounded in $[a_0, b_0]$, by the Bolzano-Weierstrass Theorem, there exists a subsequence $\{c_{k_j}\}$ that converges to some limit $c^* \in [a_0, b_0]$. Our goal is to show that $f(c^*) = 0$.
Suppose, for contradiction, that $f(c^*) \neq 0$.
Case 1: $f(c^*) > 0$:
Since f is continuous, there exists $\delta > 0$ such that $f(x) > 0$ for all $x \in (c^* - \delta, c^* + \delta)$.
Because $c_{k_j} \to c^*$, for sufficiently large j, $c_{k_j} \in (c^* - \delta, c^* + \delta)$, and $f(c_{k_j}) > 0$.
Similarly, since $f(a_{k_j}) \cdot f(b_{k_j}) < 0$, and c_{k_j} replaces either a_{k_j} or b_{k_j}, we have:

- If $f(c_{k_j}) \cdot f(a_{k_j}) < 0$, then $f(a_{k_j}) < 0$.
- If $f(c_{k_j}) \cdot f(b_{k_j}) < 0$, then $f(b_{k_j}) < 0$.

Thus, there is a sequence of points $\{d_{k_j}\} \subset [a_0, b_0]$ (either a_{k_j} or b_{k_j}) such that $f(d_{k_j}) < 0$.
Since $\{d_{k_j}\}$ is bounded in $[a_0, b_0]$, it has a convergent subsequence $\{d_{k_{j_m}}\}$ converging to some $d^* \in [a_0, b_0]$.
Because f is continuous, $f(d^*) = \lim_{m\to\infty} f(d_{k_{j_m}}) \leq 0$. In fact, $f(d^*) < 0$, since $f(d_{k_{j_m}}) < 0$ for all m.
But since $c_{k_{j_m}} \to c^*$ and $d_{k_{j_m}} \to d^*$, and $f(c^*) > 0$ and $f(d^*) < 0$, we have $f(c^*) \cdot f(d^*) < 0$.
Therefore, by the Intermediate Value Theorem, there exists a "new" point x^* (i.e. a root) between c^* and d^* such that $f(x^*) = 0$, contradicting the assumption.
Case 2: $f(c^*) < 0$:
A similar argument leads to a contradiction as well in this case, showing that must $f(c^*) = 0$.
We have shown that any convergent subsequence of $\{c_k\}$ converges to a root c^* of f. To show that $\{c_k\}$ itself converges to c^*, suppose, for contradiction, that $\{c_k\}$ has another subsequence $\{c_{k_i}\}$ converging to a different limit $c^\dagger \neq c^*$. Using the same argument as in Step 3, we can show that $f(c^\dagger) = 0$.
Since f is continuous and the initial interval $[a_0, b_0]$ contains only one root (due to the IVT and the sign change), it must be that $c^\dagger = c^*$. Therefore, all convergent subsequences of $\{c_k\}$ converge to c^*. Thus, the sequence $\{c_k\}$ converges to c^*.

□

3. Suppose $f(x)$ and $g(x)$ are functions with a common root $x = a$.

a) Prove that a solution to the homotopy continuation initial value problem

$$x'(t) = -\frac{H_t}{H_x} \qquad x(0) = a$$

is the constant function $x = a$.

b) Give an example where the solution above is *not* unique.

Hint: see handout for a picture of (a). Think about how it could be adapted (b); you can even use the tool to help you construct an example.

Solution.

Proof. **Part (a): Proving $x(t) = a$ is a Solution**
Let us define the homotopy $H(x, t)$ as

$$H(x, t) = (1 - t)f(x) + tg(x).$$

Since $f(a) = g(a) = 0$, it follows that $H(a, t) = 0$ for all $t \in [0, 1]$.
We need to show that $x(t) = a$ satisfies the differential equation

$$x'(t) = -\frac{H_t}{H_x}, \quad x(0) = a.$$

Computing the Partial Derivatives:
First, compute H_t and H_x:

$$H_t(x, t) = -f(x) + g(x),$$
$$H_x(x, t) = (1 - t)f'(x) + tg'(x).$$

Evaluate these at $x = a$:

$$H_t(a, t) = -f(a) + g(a) = -0 + 0 = 0.$$
$$H_x(a, t) = (1 - t)f'(a) + tg'(a).$$

Note that $H_x(a, t)$ may not be zero unless both $f'(a)$ and $g'(a)$ are zero.
Computing $x'(t)$ at $x = a$:
Substitute $x(t) = a$ into the differential equation:

$$x'(t) = -\frac{H_t(a, t)}{H_x(a, t)} = -\frac{0}{H_x(a, t)} = 0.$$

Therefore,

$$x'(t) = 0, \quad x(0) = a.$$

This implies that $x(t) = a$ for all $t \in [0, 1]$.
Conclusion:
The constant function $x(t) = a$ is a solution to the homotopy continuation initial value problem. ☐

Proof. **Part (b): Example Where the Solution is Not Unique**
We will construct specific functions $f(x)$ and $g(x)$ with a common root at $x = a$ such that the initial value problem

$$x'(t) = -\frac{H_t}{H_x}, \quad x(0) = a,$$

has multiple solutions.
Example Functions:
Let

$$f(x) = (x - a)^{1/3}, \quad g(x) = -(x - a)^{1/3}.$$

Both functions have a root at $x = a$:

$$f(a) = g(a) = 0.$$

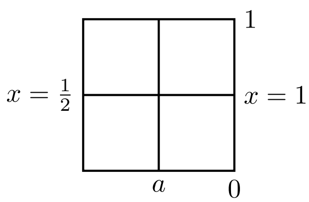
Constructing the Homotopy:
Define

$$H(x,t) = (1-t)f(x) + tg(x) = (1-t)(x-a)^{1/3} + t(-(x-a)^{1/3}) = (1-2t)(x-a)^{1/3}.$$

Compute H_t and H_x:

$$H_t(x,t) = -f(x) + g(x) = -(x-a)^{1/3} - (x-a)^{1/3} = -2(x-a)^{1/3},$$

$$H_x(x,t) = (1-2t) \cdot \frac{1}{3}(x-a)^{-2/3}.$$



When t shifts from $0 \to 1$, at a certain point for t the homotopy will coincide with the x-axis completely. This leads a shift from the solution $x = a$ when $t = 0$ to the solution x equals to every possible points on the x-axis between $x = \frac{1}{2}$ to $x = 1$. This clearly makes the solution **NOT Unique** during the approximation.

$\square$