Homework 6 (root finding)

Hanzhang Yin

November 18, 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. We proceed in several steps to establish the convergence of $\{c_k\}$ to a root of f.
Given a_k and b_k such that $f(a_k) \cdot f(b_k) < 0$, define

$$c_k = b_k - f(b_k) \cdot \frac{b_k - a_k}{f(b_k) - f(a_k)}.$$

This is the point where the secant line through $(a_k, f(a_k))$ and $(b_k, f(b_k))$ crosses the x-axis.
Now, we claim that:

1. The sequence of intervals $[a_k, b_k]$ is nested, i.e.,

$$[a_{k+1}, b_{k+1}] \subseteq [a_k, b_k], \quad \forall k \geq 0.$$

2. The function f changes sign on each $[a_k, b_k]$, i.e.,

$$f(a_k) \cdot f(b_k) < 0, \quad \forall k \geq 0.$$

Justification:
At each iteration, we select a_{k+1} and b_{k+1} such that one of them is c_k and the other is either a_k or b_k, depending on the sign of $f(c_k)$. Specifically:
If $f(a_k) \cdot f(c_k) < 0$, set $a_{k+1} = a_k$ and $b_{k+1} = c_k$.
If $f(c_k) \cdot f(b_k) < 0$, set $a_{k+1} = c_k$ and $b_{k+1} = b_k$.
In either case, $[a_{k+1}, b_{k+1}] \subseteq [a_k, b_k]$ and f changes sign on $[a_{k+1}, b_{k+1}]$.
Since $\{[a_k, b_k]\}$ is a nested sequence of closed intervals with $a_k \leq b_k$, and f is continuous on $[a_0, b_0]$, we can consider the lengths $\ell_k = b_k - a_k$. We need to show that $\ell_k \to 0$ as $k \to \infty$.
Let $I = \bigcap_{k=0}^{\infty} [a_k, b_k]$. Since the intervals are nested and closed, and $a_k \leq b_k$, the intersection I is non-empty and closed by "Nested Interval Convergence Theorem"
Assuming $\ell_k \to 0$, it follows that I contains exactly one point, say c^*. Therefore,

$$\lim_{k\to\infty} a_k = \lim_{k\to\infty} b_k = c^*.$$

Since $c_k \in [a_k, b_k]$, we also have

$$\lim_{k\to\infty} c_k = c^*.$$

Because f is continuous on $[a_0, b_0]$, and $f(a_k) \cdot f(b_k) < 0$ for all k, we have

$$f(a_k) \to f(c^*), \quad f(b_k) \to f(c^*).$$

But since $f(a_k)$ and $f(b_k)$ have opposite signs for each k, it must be that $f(c^*) = 0$. Otherwise, $f(a_k)$ and $f(b_k)$ would eventually have the same sign, contradicting $f(a_k) \cdot f(b_k) < 0$. Therefore, c^* is a root of f.

Unlike the bisection method, the lengths ℓ_k do not necessarily decrease by a fixed ratio. However, we still can show that $\ell_k \to 0$ under the assumption that f is continuous on $[a, b]$.
Suppose, for contradiction, that ℓ_k does not converge to zero. Then there exists $\epsilon > 0$ such that, $\ell_k \geq \epsilon$. This would imply that there is an infinite number of intervals $[a_k, b_k]$ with length at least ϵ. Since $[a_k, b_k] \subseteq [a_0, b_0]$, the nested intervals would not converge to a single point, contradicting the fact that the intersection I contains exactly one point c^*. Therefore, $\ell_k \to 0$. □

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:

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

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

Observations:
At $x = a$, the term $(x-a)^{-2/3}$ becomes undefined because of division by zero. Thus, $H_x(a,t)$ is not defined, and the differential equation involves an expression of the form $0/0$ at $x = a$.
Analyzing the Differential Equation:
The differential equation becomes

$$x'(t) = -\frac{H_t}{H_x} = -\frac{-2(x-a)^{1/3}}{(1-2t) \cdot \frac{1}{3}(x-a)^{-2/3}} = \frac{-2(x-a)^{1/3}}{(1-2t) \cdot \frac{1}{3}(x-a)^{-2/3}}.$$

Simplify the right-hand side:

$$x'(t) = \frac{-2(x-a)^{1/3}}{(1-2t) \cdot \frac{1}{3}(x-a)^{-2/3}} = -6 \cdot \frac{(x-a)^{1/3}}{(1-2t)(x-a)^{-2/3}} = -6 \cdot \frac{(x-a)^{1/3+2/3}}{1-2t} = -6 \cdot \frac{x-a}{1-2t}.$$

Resulting Differential Equation:
We obtain

$$x'(t) = -6 \cdot \frac{x-a}{1-2t}.$$

Solving the Differential Equation:
Separate variables and integrate:

$$\frac{x'(t)}{x-a} = -6 \cdot \frac{1}{1-2t}.$$

Let $y(t) = x(t) - a$. Then

$$\frac{y'(t)}{y(t)} = -6 \cdot \frac{1}{1-2t}.$$

Integrate both sides:

$$\int \frac{y'(t)}{y(t)}\, dt = -6 \int \frac{1}{1-2t}\, dt,$$

$$\ln|y(t)| = 3\ln|1-2t| + C,$$

$$\ln|y(t)| = \ln|(1-2t)^3| + C.$$

Exponentiate both sides:

$$|y(t)| = e^C |(1-2t)^3|.$$

Since $y(0) = x(0) - a = 0$, we have $|y(0)| = 0$, which implies that $e^C |(1-0)^3| = 0$. This is only possible if $e^C = 0$, which is impossible.
Therefore, the only solution satisfying $x(0) = a$ is $y(t) = 0$, i.e.,

$$x(t) = a.$$

However, due to the singularity at $x = a$, the uniqueness theorem for differential equations (Picard-Lindelöf) does not apply because the right-hand side is not Lipschitz continuous at $x = a$. This allows for the existence of multiple solutions. $\square$