Convergence of Secant/Newton's Method in 1D

Mei Shin Lee

February 2nd, 2022

1 Minimizing a Function using Newton's Method

For solving equations $f(x) = 0$ we iterated using

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

Let's say we want to solve

$$x^* = \arg\min_{a \leq x \leq b} f(x)$$

If f is $C'([a, b])$ (continuous, differentiable on (a, b)) and its derivative f' is continuous, then a "first order necessary condition for optimality" is $f'(x^*) = 0$
In detail,

- "first order": Look at f'
- "necessary condition": Must be the interior point minimum. While this statement may be true, it doesn't always imply that it is a minimum.
- "optimality": minimum

A second order sufficient condition for minimization would be $f''(x^*) > 0$.
Lastly, to determine the behavior of the x^* at the endpoints, we will use Lagrange Multipliers.

What if we apply Newton's Method to the 1st order necessary condition?
Set $g(x) = f'(x)$.

$$x_{n+1} = x_n - \frac{g(x_n)}{g'(x_n)} = x_n - \frac{f'(x_n)}{f^*(x_n)}$$

What happens if $f(x) = ax^2 + bx + c$?

$$\begin{aligned}
x_0 &= x\ldots \text{ (doesn't matter)} \\
f'(x) &= 2ax + b \\
f''(x) &= 2a \\
x_1 &= x_0 - \frac{2ax + b}{2a} = \frac{-b}{2a}
\end{aligned}$$

Now if $\frac{-b}{2a}$ is substituted into $f'(x)$, observe that

$$f'(\frac{-b}{2a}) = 2a(\frac{-b}{2a}) + b = 0$$

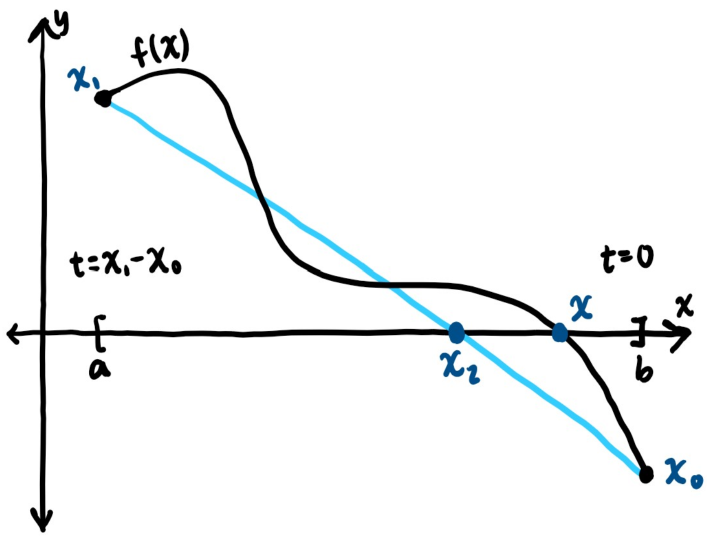
It converges in one step.

Exercise: Show that minimizing a function $f : \mathbb{R} \to \mathbb{R}$ using Newton's method is equivalent to minimizing a sequence of quadratics.
Hint: Do a $k = 2$ Taylor Expansion of $f(x_n + \Delta x_n)$ about x_n where $\Delta x_n = x_{n+1} - x_n$

Bonus Exercise: Apply the above exercise to the Secant Method.

2 Convergence of the Secant Method

Recall the Secant Method, where we solve for $f(x) = 0$



$$\begin{aligned}
l(t) &= \frac{f_1 - f_0}{x_1 - x_0}(t) + f_0 \\
l(0) &= f_0 \\
l(x_1 - x_0) &= \frac{f_1 - f_0}{x_1 - x_0} + f_0 \\
&= f_1 - f_0 + f_0 \\
&= f_1
\end{aligned}$$

Now, find t such that $l(t) = 0$

$$\begin{aligned} t &= -f_0 \cdot \frac{x_1 - x_0}{f_1 - f_0} \\ x_2 - x_1 &= -f_0 \cdot \frac{x_1 - x_0}{f_1 - f_0} \\ x_2 &= x_1 - f_0 \cdot \frac{x_1 - x_0}{f_1 - f_0} \\ x_{n+1} &= x_n - f_n \cdot \frac{x_n - x_{n-1}}{f_n - f_{n-1}} \end{aligned}$$

The derived equation: $x_{n+1} = x_n - f_n \cdot \frac{x_n - x_{n-1}}{f_n - f_{n-1}}$ is the secant iteration.

How do we address the question of how fast a sequence approach its limit?

Definition: Let $\xi = \lim_{x\to\infty} x_n$. Then we say that $x_n \to \xi$ with order $q > 1$ if the sequence $\varepsilon_n = |\xi - x_n|$ (difference of current iterate and target) converges to 0 and there exists some $\mu \in (0, 1)$ such that

$$\lim_{x\to\infty} \frac{\varepsilon_{n+1}}{\varepsilon_n^q} = \mu$$

If $q = 1$, we say that it converges linearly. If $q = 2$, it converges quadratically. Newton's method converges with order 2.

Theorem: If the secant method converges, then it converges with the rate $q = \frac{1+\sqrt{5}}{2}$.
Proof: Assume $x_n \to \xi$ and $\varepsilon_n = x_n - \xi$
We have the secant iteration

$$x_{n+1} = x_n - f_n \cdot \frac{x_n - x_{n-1}}{f_n - f_{n-1}}$$

Subtract ξ from both sides:

$$\begin{aligned} \varepsilon_{n+1} &= \varepsilon_n - f_n \cdot \frac{x_n - x_{n-1}}{f_n - f_{n-1}} \\ &= \varepsilon_n - f_n \cdot \frac{\varepsilon_n - \varepsilon_{n-1}}{f_n - f_{n-1}} \\ &= \frac{\varepsilon_n(f_n - f_{n-1}) - f_n(\varepsilon_n - \varepsilon_{n-1})}{f_n - f_{n-1}} \\ &= \frac{f_n\varepsilon_{n-1} - \varepsilon_n f_{n-1}}{f_n - f_{n-1}} \end{aligned}$$

Remember: $x_n = \varepsilon_n + \xi$ and f_n is just notation for $f(x_n)$. So we Taylor expand about ξ.

$$\begin{aligned} f_n &= f(x_n) = f(\varepsilon + \xi) \\ &= f(\xi) + \varepsilon_n f'(\xi) + \varepsilon^2_{\frac{n}{2}} f''(\xi) + O(\varepsilon_n^3) \\ &= 0 + \varepsilon_n f'(\xi) + \varepsilon^2_{\frac{n}{2}} f''(\xi) + O(\varepsilon_n^3) \end{aligned}$$

Write $f^{(p)}(\xi) = f_*^{(p)}$. (So, $f(\xi) = f_*, f'(\xi) = f'_*$..., etc.)
Taylor expand f_n (done above) and f_{n-1} about ξ to get:

$$f_n = \varepsilon_n f'_* + \frac{\varepsilon_n^2}{2} f''_* + O(\varepsilon_n^3)$$
$$f_{n-1} = \varepsilon_{n-1} f'_* + \frac{\varepsilon_{n-1}^2}{2} f''_* + O(\varepsilon_{n-1}^3)$$

Now substitute: $f'_* = f'(\xi)$

$$\begin{aligned}
\varepsilon_{n+1} &= \frac{[(\varepsilon_n f'_* + \frac{\varepsilon_n^2}{2} f''_* + O(\varepsilon_n^3))\varepsilon_{n-1} - (\varepsilon_{n-1} f'_* + \frac{\varepsilon_{n-1}^2}{2} f''_* + O(\varepsilon_{n-1}^3))\varepsilon_n]}{\varepsilon_n f'_* + \frac{\varepsilon_n^2}{2} f''_* + O(\varepsilon_n^3) - \varepsilon_{n-1} f''_* - \frac{\varepsilon_{n-1}^2}{2} f''_* + O(\varepsilon_{n-1}^3)} \\
&= \frac{\varepsilon_n \varepsilon_{n-1} [\frac{\varepsilon_n - \varepsilon_{n-1}}{2} f''_* + O(\varepsilon_n^2) + O(\varepsilon_{n-1}^2)]}{(\varepsilon_n - \varepsilon_{n-1}) f'_* + O(\varepsilon_n^2) + O(\varepsilon_{n-1}^2)} \\
\lim_{n\to\infty} \frac{\varepsilon_{n+1}}{\varepsilon_n \varepsilon_{n-1}} &= \lim_{n\to\infty} \frac{[\frac{\varepsilon_n \varepsilon_{n-1}}{2} f''_* + O(\varepsilon_n^2)]}{[(\varepsilon_n \varepsilon_{n-1}) f'_* + O(\varepsilon_n^2)]} \\
&= \frac{f''_*}{2 f'_*}
\end{aligned}$$

We need to find a constant $\mu > 0$ such that $\frac{\varepsilon_{n+1}}{\varepsilon_n^2} \to \mu$ as $n \to \infty$
Let's just say that $\varepsilon_{n+1} = C\varepsilon_n^q$

$$\begin{aligned}
\frac{\varepsilon_{n+1}}{\varepsilon_n \varepsilon_{n-1}} &= \frac{C\varepsilon_n^q}{\varepsilon_n \varepsilon_{n-1}} \\
&= \frac{C[C\varepsilon_{n-1}^q]^q}{C\varepsilon_{n-1}^q \varepsilon_{n-1}} \\
&= \frac{C \cdot C^q \cdot q_{n-1}^{q^2}}{C\varepsilon_{n-1}^{q+1}} \\
&= C^q \varepsilon_{n-1}^{q^2-q-1} \\
\lim_{n\to\infty} \frac{\varepsilon_{n+1}}{\varepsilon_n \varepsilon_{n-1}} &= \lim_{n\to\infty} C^q \varepsilon_{n-1}^{q^2-q-1}
\end{aligned}$$

Now we pick q such that $q^2 - q - 1 = 0$

$$q = \frac{1 \pm \sqrt{1+4}}{2} = \frac{1 \pm \sqrt{5}}{2} < 0$$

Choose

$$q = \frac{1+\sqrt{5}}{2}$$

This argument shows that if the Secant Method converges, it does so with order $q = \frac{1+\sqrt{5}}{2}$.

3 A Common Pattern with Taylor Expansions

1. Do two similar Taylor Expansions
2. Look for cancellations with every other power

Example:

$$\begin{aligned} F(x+h) &= F(x) + hF'(x) + \frac{h^2}{2}F''(x) + O(h^3) \\ F(x-h) &= F(x) - hF'(x) + \frac{h^2}{2}F''(x) + O(h^3) \\ F(x+h) - F(x-h) &= 2hF'(x) + O(h^3) \end{aligned}$$

Rearrange:

$$F'(x) = \frac{F(x+h) - F(x-h)}{2h} + O(h^2)$$

The result looks similar to $F'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$. This is called a *finite difference approximation.*

4 Taylor Expansion Remainders

Another form is the so-called Lagrange form of the remainder. It looks like:

$$f(x+h) = \sum_{m=0}^{k} \frac{f^{(m)}}{m!}h^m + \frac{f^{(m+1)}\eta}{(m+1)!}h^{m+1}, \eta \in [0,h]$$

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + ... + \frac{f^{(m)}x_0}{m!}(x-x_0)^m + \frac{f^{(m+1)}\eta}{(m+1)!}(x-x_0)^{m+1}, \eta \in [x, x_0]$$

$$\begin{aligned} 0 = f(\xi) &= f(x_n + \xi - x_n) \\ &= f(x_n) + f'(x_n)(\xi - x_n) + \frac{f''(\eta)}{2}(\xi - x_n)^2 \end{aligned}$$

Rearrange and apply the Newton step to get:

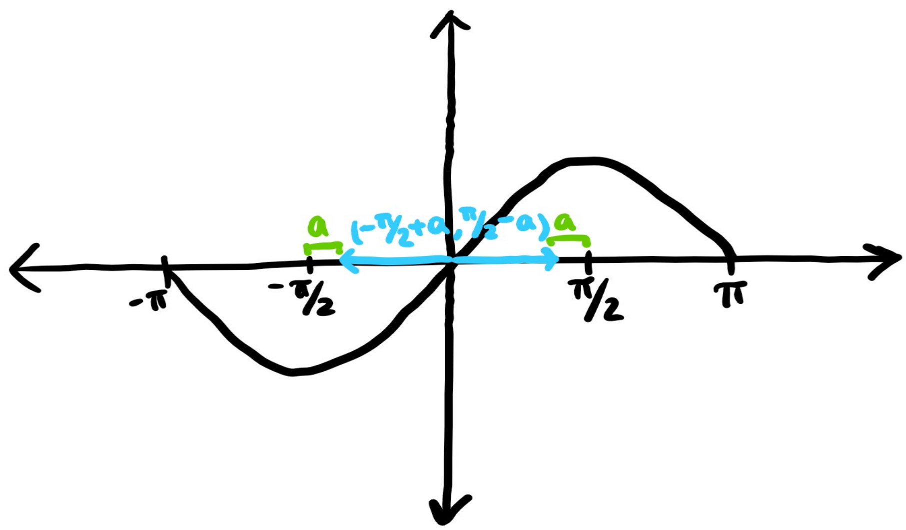
$$\begin{aligned} \xi - x_{n+1} &= -\frac{(\xi - x_n)^2}{2} \cdot \frac{f''(\eta)}{f'(x_n)} \\ \lim_{n\to\infty} \frac{\varepsilon_{n+1}}{\varepsilon_n^2} &= \frac{-1}{2} \cdot \frac{f''(\xi)}{f'(\xi)} \\ &\to q = 2 \text{ for Newton} \end{aligned}$$

5 Theorem for the Homework Problem (Q2)

Note: This ended up not being given for Written Homework #1 as per Professor Potter's email.

Theorem: Let $f \in C^2$ on $I_\delta = [\xi - \delta, \xi + \delta], \delta > 0$. Assume that $f(\xi) = 0$ and $f''(\xi) \neq 0$. Assume that $\exists A > 0$ such that $\frac{|f''(x)|}{|f'(y)|} \leq A$ for all $x, y \in I_\delta$. If x_0 is such that $|\xi - x_0| \leq min(\delta, \frac{1}{A})$, then $x_n \to \xi$ quadratically.

Exercise (HW): Let $f(x) = sin(x)$, so $sin(0) = 0$.



Apply this theorem to show that if $x_0 \in (\frac{-\pi}{2} + a, \frac{\pi}{2} - a)$, where $a \geq 0$ and $x_0 \neq 0$, $x_n \to 0$ quadratically.
Note: Do you have to assume anything about a?