Analysis of Newton's method

What you should learn:

- how to quantify "speed" of convergence
- how to prove Newton converges for x_0 "close enough" to x^*, or for any x_0 if $f(x)$ is convex.

Definition: order of convergence

Suppose $x_n \to \xi$. x_n converges to ξ at order p if

$$\lim_{k \to \infty} \frac{|x_{k+1} - \xi|}{|x_k - \xi|^p} = M$$

where $M > 0$ is independent of k.

Note: if $p=1$, need $M<1$ for convergence.

Recall Taylor w/ remainder

Lemma: Suppose $f \in C^{n+1}[a,b]$

$\Rightarrow$ for all $x \in [a,b]$, $\exists \eta \in [a,b]$

s.t. $f(x) = f(a) + f'(a)(x-a) + \cdots$

$$+ \frac{f^{(n+1)}(\eta)}{(n+1)!}(x-a)^{n+1}$$

Error equation for Newton

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

$$x_n \to \xi \Rightarrow \underbrace{(x_{k+1} - \xi)}_{e_{k+1}} = \underbrace{(x_k - \xi)}_{e_k} - \frac{f(x_k)}{f'(x_k)}$$

$$\Rightarrow e_{k+1} = e_k - \frac{f(x_k)}{f'(x_k)}$$

(Taylor) $f(\xi) = 0 = f(x_k) - f'(x_k)\underbrace{(x_k - \xi)}_{e_k}$

(if $f \in C^2[x_k, \xi]$) $+ \frac{1}{2} f''(\eta_k)\underbrace{(x_k-\xi)^2}_{e_k^2}$

$$f(x_k) = f'(x_k)e_k - \frac{1}{2}f''(\eta_k)\, e_k^2$$

Sub into error eqn.

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

$$e_{k+1} = e_k - \frac{f'(x_k)\, e_k - \frac{1}{2} f''(\eta_k) e_k^2}{f'(x_k)}$$

$$e_{k+1} = \underbrace{\frac{\frac{1}{2} f''(\eta_k)}{f'(x_k)}}\, e_k^2$$

$$|e_{k+1}| \leq C_k e_k^2 \Rightarrow \text{quadratic convergence if } |C_k| \leq C.$$

<u>More rigorous analysis:</u>

<u>Thm:</u> Let $f(\xi) = 0$ & let $f \in C^2(I_\delta)$ w/ $I_\delta = [\xi - \delta, \xi + \delta]$ for some $\delta > 0$.

Assume $f'(\xi), f''(\xi) \neq 0 \Rightarrow \xi$ root has multiplicity 1

& $\left|\frac{f''(x)}{f'(y)}\right| \leq A \quad \forall\, x, y \in I_\delta$

Then, if $|x_0 - \xi| \leq h = \min\{\delta, \frac{1}{A}\}$

then $x_k \to \xi$ quadratically.

Proof: ① Just show Newton converges

② Show Newton converges quadratically.

inductively: x_k conditions $\Rightarrow$ x_{k+1} condition

Recall Taylor: $0 = f(\xi) = f(x_k) + f'(x_k)(\xi - x_k) + \frac{f''(\eta_k)}{2}(\xi - x_k)^2$

(w/ η_k b/w ξ, x_k)

$\Rightarrow$ inductive assumption $|x_k - \xi| \leq h \leq \delta$

Recall error eqn: $|e_{k+1}| \leq \left|\frac{1}{2}\frac{f''(\eta_k)}{f'(x_k)}\right| |e_k^2|$

$$\leq \frac{A}{2}|e_k^2|$$

But we've also assumed $\underbrace{|x_k - \xi|}_{|e_k|} \leq h \leq \frac{1}{A}$

$\Rightarrow |e_{k+1}| \leq \frac{1}{2} \cancel{A} |e_k| \frac{1}{\cancel{A}} = \frac{1}{2}|e_k|$

$\Rightarrow$ error decrease w/ k $\Rightarrow$

Newton converges (linearly)

2

$$\frac{|e_{k+1}|}{|e_k|^2} \le \left| \frac{1}{2} \frac{f''(\eta_k)}{f'(x_k)} \right|$$

$f \in C^2(I_\delta)$

$\Rightarrow$ as $x_k \to \xi$

$\eta_k \in [x_k, \xi]$

so $\eta_k \to \xi$

$\Downarrow$

converges to $\frac{1}{2} \left| \frac{f''(\xi)}{f'(\xi)} \right|$

$$\frac{|e_{k+1}|}{|e_k|^2} \le \frac{1}{2} \left| \frac{f''(\xi)}{f'(\xi)} \right| > 0$$

Needed $|x_0 - \xi| \le h \le \min\{\delta, \frac{1}{A}\}$

$\delta \Rightarrow I_\delta$ over which $f \in C^2(I_\delta)$

$A \Rightarrow f'', f'$

$\Rightarrow$ can relax condition if f is convex

Thm: (Suli/Mayer 1.9). Suppose f is $C^2(I_\delta)$, $f''(\xi), f'(\xi) \neq 0$.

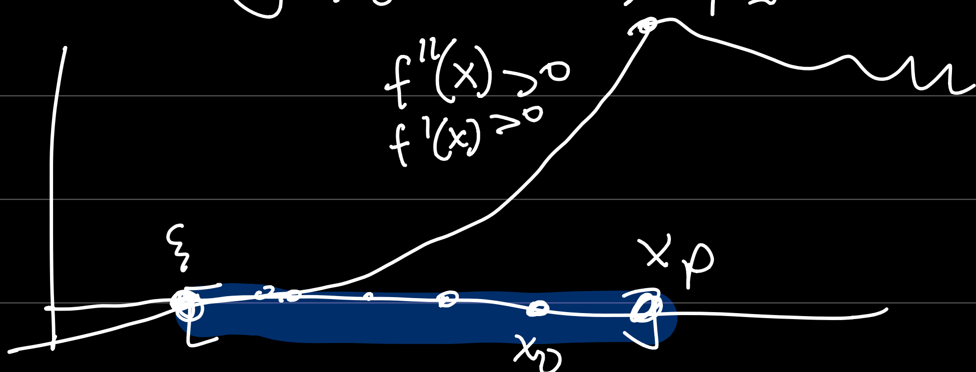
Then, if we assume $\exists\, x_p \in \mathbb{R}$ s.t.

$x_p > \xi$ &

$f'(x), f''(x) > 0$ on $[\xi, x_p]$

$\Rightarrow$ Newton converges (quadratically) for any $x_0 \in [\xi, x_p]$

Proof



If we show convergence (linearly) then we can reuse ② of previous proof.

$\Rightarrow$ Let $x_k \in [\xi, x_p]$, then

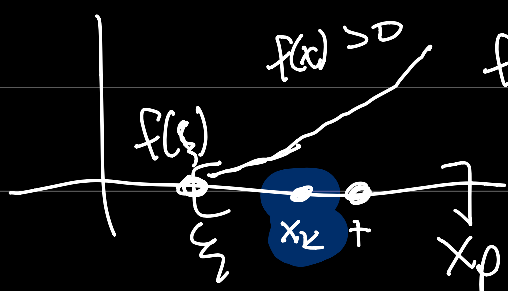
$$\underbrace{e_{k+1}}\ = e_k^2 \frac{f''(\eta_k)}{f'(x_k)} \Rightarrow \eta_k \in [\xi, x_k]$$

$$x_{k+1} = \xi + e_k^2 \frac{f''(\eta_k)}{f'(x_k)}$$

$$x_{k+1} > \xi$$

However, note $f'(x) > 0$ on $[\xi, x_p]$

so b/c $f(\xi) = 0$



$$f(x) = \int_{\xi}^{x} f'(s)\, ds$$

$$\Rightarrow f(x) > 0$$

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

$f(x_k) > 0$

$f'(x_k) > 0$

$x_{k+1} = x_k -$ something pos

$$\underbrace{\xi < x_{k+1}}_{\text{prev. part.}} < x_k$$

$\Rightarrow \lim_{k \to \infty} x_k \to \xi$

$\underbrace{\text{Newton}}_{\text{requires } f'(x)}$, bisection

requires $f'(x) \Longrightarrow$
- explicitly
- Automatic differentiation
- method to approx $f'(x)$

Is there a method faster than bisection w/out needing $f'(x)$?

Yes $\Rightarrow$ secant (+ more).

Motivation: $\lim_{h \to 0} f'(x) = \dfrac{f(x+h) - f(x)}{h}$

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

$x = x_{k-1}$

$h = x_k - x_{k-1}$

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