(proof of convergence of bisection method; slides p.6)

Chop the task into pieces.

① $\lim a_n$, $\lim b_n$, $\lim c_n$ exist and they all the same.

② Call the limit, ξ, then $f(\xi)=0$.

③ $|c_n - \xi| < 2^{-(n+1)}(b-a)$

---

① Observe $\underbrace{a_0 \le a_1 \le a_2 \le \cdots}_{\text{nondecreasing}} \underbrace{\le b}_{\text{bdd above}}$ by construction. Therefore $\lim a_n$ exists.

/* Math 3B – monotone sequence theorem.

If $\{a_n\}$ is nondecreasing (or nonincreasing) and bounded above (or bounded below), the limit exists. [This is "half-version"]

If $\{a_n\}$ is monotonic (i.e., only nonincreasing or only nondecreasing) and bounded (i.e., bounded from above and below), it converges. [This is "two-sided-ver"] */

Likewise $b_0 \ge b_1 \ge b_2 \ge \cdots \ge a$. Therefore b_n also converges.

Let $\lim a_n = \xi_1$ and $\lim b_n = \xi_2$.
We know the length of $[a_n, b_n]$ gets halved from the construction. Therefore, $b_n - a_n \to 0$ as $n \to \infty$. Then, we must have

$$0 = \lim(b_n - a_n) = \lim b_n - \lim a_n$$
$$= \xi_2 - \xi_1$$

$\Longrightarrow \xi_1 = \xi_2$ Call this common limit ξ.
Lastly for ①, we have $a_n \leq c_n \leq b_n$.
Therefore, sandwich theorem says

$$\underset{\substack{\downarrow \\ \xi}}{\lim a_n} \leq \underbrace{\lim c_n}_{\substack{\downarrow \\ \xi}} \leq \underset{\substack{\downarrow \\ \xi}}{\lim b_n}$$

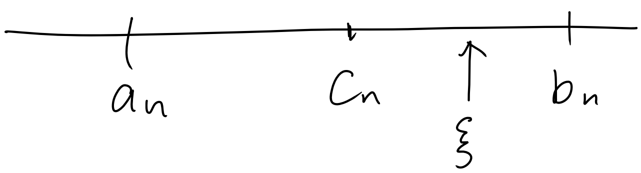
Thus, $\lim c_n = \xi$.

② Taking limit of $f(a_n)f(b_n) \le 0$, we have $f(\xi)f(\xi) \le 0$. The only possibility is $f(\xi) = 0$.

/* Here, continuity of f is used.

f conti at a $\iff$ $f(a_n) \to f(a)$ $\forall \{a_n\}$ s.t. $a_n \to a$

③ Observe $\xi \in [a_n, b_n]$ $\forall n$. Thus,



$$
\begin{aligned}
|c_n - \xi| &< \frac{1}{2} \cdot \text{length}([a_n, b_n]) \\
&= \frac{1}{2} \cdot \frac{1}{2} \text{length}([a_{n-1}, b_{n-1}]) \\
&\ \ \vdots \\
&= \left(\frac{1}{2}\right)^n \text{length}([a_1, b_1]) \\
&= \left(\frac{1}{2}\right)^{n+1} \underbrace{\text{length}([a_0, b_0])}_{= b_0 - a_0 = b - a}
\end{aligned}
$$

(proof of quadratic conv. of Newton's method.)

Put $e_n = x_n - \xi$.

Subtract ξ from the method and sneak in $f(\xi)$

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

(*)

$$e_{n+1} = \frac{1}{f'(x_n)}\left(f(\xi) - f(x_n) - f'(x_n)(\xi - x_n) \right)$$

$$\left(f(\xi) = f(x_n) + f'(x_n)(\xi - x_n) + \frac{f''(c_n)}{2!}(\xi - x_n)^2, \quad c_n \in (\xi, x_n) \text{ or } (x_n, \xi) \right)$$

$$= \frac{1}{f'(x_n)} \cdot \frac{f''(c_n)}{2} e_n^2$$

as $e_n \to 0$

Therefore, IF $e_n \to 0$ as $n \to \infty$, then $x_n \to \xi$, and in turn, $f'(x_n) \to f'(\xi)$ and $f''(c_n) \to f''(\xi)$ as $n \to \infty$.

Then, dividing the error eqn (*) by e_n^2 and taking limit $e_n \to 0$, we have

$$\frac{e_{n+1}}{e_n^2} \longrightarrow \frac{f''(\xi)}{2f'(\xi)} \text{ as } e_n \to 0.$$

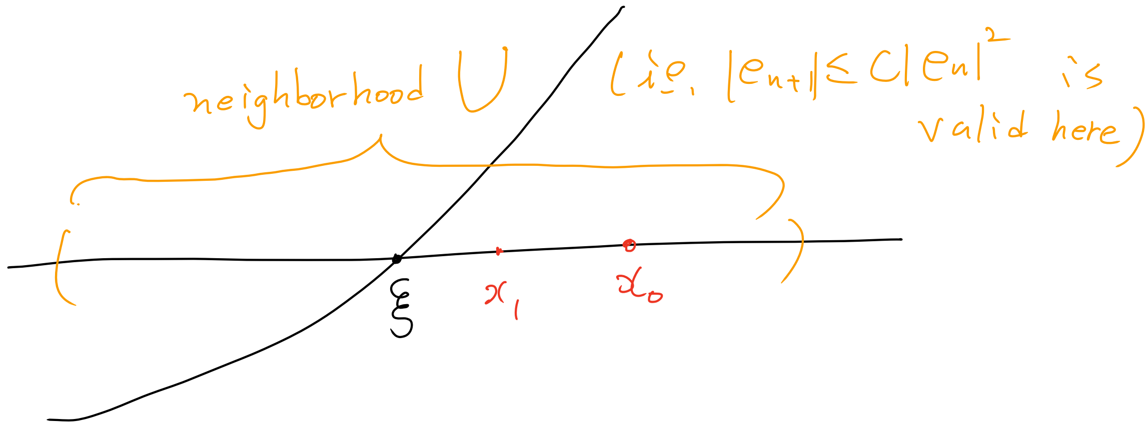
That is,

$$\left| \frac{e_{n+1}}{e_n^2} \right| \lesssim C$$

if $e_n \simeq 0$. Or roughly "$|e_{n+1}| \lesssim C|e_n|^2$"

↳ just a generic constant as long as it is fixed.

Now, we prove the IFF part.



Let $|e_{n+1}| \le C|e_n|^2$ on U

Choose x_0 so that $|e_0| = |x_0 - \xi| < \frac{1}{2C}$ and $x_0 \in U$.

/* This is where "if the initial guess is sufficiently close to the zero" comes in. */

Then, we see

$$|e_1| \le C|e_0|^2 \le C \cdot \frac{1}{2C} \cdot |e_0| = \frac{1}{2}|e_0|$$

Therefore, $x_1 \in U$ too.

Repeat this so that

$$|e_2| \le C|e_1|^2 \le C \cdot \left(\frac{1}{2}\right)^2 |e_0| \cdot |e_0|$$
$$\le C \cdot \left(\frac{1}{2}\right)^2 \frac{1}{2C} \cdot |e_0| = \left(\frac{1}{2}\right)^3 |e_0|$$

$$\cdots \quad |e_n| \le \left(\frac{1}{2}\right)^{2n-1} |e_0|$$

Thus, $e_n \to 0$ as $n \to \infty$. ▨

$\boxed{\text{Thm}}$ (Taylor thm)

If $f \in C^{n+1}$ near a point x (i.e. on $(x-\delta, x+\delta)$),
then for any $y \in (x-\delta, x+\delta)$, we have

(Lagrange remainder ver.)

$$f(y) = f(x) + f'(x)(y-x) + \frac{f''(x)}{2!}(y-x)^2 + \cdots + \frac{f^{(n)}(x)}{n!}(y-x)^n + \frac{f^{(n+1)}(\xi_y)}{(n+1)!}(y-x)^{n+1} \quad (\xi_y \in (x,y) \text{ or } (y,x))$$

and

(Integral remainder ver.)

$$f(y) = f(x) + f'(x)(y-x) + \frac{f''(x)}{2!}(y-x)^2 + \cdots + \frac{f^{(n)}(x)}{n!}(y-x)^n + \int_x^y \frac{f^{(n+1)}(t)}{n!} \cdot (y-t)^n \, dt$$

(proof of Taylor theorem with Lagrange remainder)

If $y=x$, there is nothing to prove. So, $y \neq x$.

Set $M = \left(f(y) - T_n(y;x)\right) / (y-x)^{n+1}$

so that $f(y) - T_n(y;x) - M(y-x)^{n+1} = 0$.

We want: there is $\xi_y \in (x,y)$ s.t. $f^{(n+1)}(\xi_y) = (n+1)!M$

Introduce

$$g(t) = f(t) - T_n(t;x) - M(t-x)^{n+1}$$

Note that $f(x) = T_n(x;x)$, $f'(x) = T_n'(x;x)$, $\ldots$, $f^{(n)}(x) = T_n^{(n)}(x;x)$. Therefore,

$g(x) = g'(x) = g''(x) = \cdots g^{(n)}(x) = 0$ since $M(t-x)^{n+1}$ has zero at x of order $n+1$.

By construction of M, we have $g(y)=0$

Apply MVT (mean value theorem) to g. Then, there is $\xi_1 \in (x,y)$ s.t. $g'(\xi_1)=0$. Apply MVT to g' on $[x,\xi_1]$. Then, there is $\xi_2 \in (x,\xi_1)$ s.t. $g''(\xi_2)=0$. Repeat this so that there is $\xi_n \in (x,\xi_{n-1})$ s.t. $g^{(n)}(\xi_n)=0$.

Repeat once more to have $\xi_{n+1} = \xi_y \in (x,\xi_n) \subset (x,y)$ s.t. $g^{(n+1)}(\xi_y)=0$. That is $f^{(n+1)}(\xi_y) = M(n+1)!$ ∎

(proof of global conv. of Newton's method for convex fn's)

Since f is increasing and has a zero by the assumption, zero is unique.

Since $f \in C^2$ and convex, $f''(x) \geq 0$ $\forall x \in \mathbb{R}$. Also, $f'(x) > 0$ $\forall x \in \mathbb{R}$ since f is increasing.

Now, recall the error equation (*)

$$e_{n+1} = \frac{1}{f'(x_n)}\left(f(\xi) - f(x_n) - f'(x_n)(\xi - x_n) \right)$$

$$= \underbrace{\frac{1}{f'(x_n)} \cdot \frac{f''(c_n)}{2}}_{\text{positive}} e_n^2$$

Deduce $e_{n+1} = x_{n+1} - \xi \geq 0$ no matter $e_n \geq 0$ or not. $\Longrightarrow$ $e_1, e_2, e_3, \cdots \geq 0$.

This, in turn, yield $f(x_n) \geq f(\xi) = 0$ $(n=1,2,\cdots)$

since f increasing: $x_n \geq \xi \Rightarrow f(x_n) \geq f(\xi)$

From $\underbrace{x_{n+1} - \xi}_{e_{n+1}} = \underbrace{x_n - \xi}_{e_n} - \underbrace{\frac{f(x_n)}{f'(x_n)}}_{\text{positive}}$,

We have $e_{n+1} \leq e_n$ for $n = 1, 2, \cdots$

Again, monotone sequence theorem says $\{e_n\}$ converges. But we don't know the limit. Call it e.

$e_{n+1} \leq e_n$ also implies $x_n \geq x_{n+1} \geq \cdots \geq \xi$. The same theorem applies so that $\lim x_n$ exists. Call it z. Taking $n \to \infty$ in the following

$$\underbrace{x_{n+1} - \xi}_{e_{n+1}} = \underbrace{x_n - \xi}_{e_n} - \underbrace{\frac{f(x_n)}{f'(x_n)}}_{\text{positive}}$$

$$\downarrow$$

$$\cancel{e} = \cancel{e} - \frac{f(z)}{f'(z)} \Rightarrow f(z) = 0$$

That is $z = \xi$, the unique zero.

(Newton's method for $\sqrt{a}$.)

Given $a > 0$, $\sqrt{a}$ can be defined by a positive root of

$$f(x) = x^2 - a.$$

Thus,

$$f'(x) = 2x.$$

Therefore the Newton's method reads

$$\begin{aligned} x_{n+1} &= x_n - f(x_n)/f'(x_n) \\ &= x_n - (x_n^2 - a)/2x_n \\ &= x_n - \frac{x_n}{2} + \frac{a}{2x_n} \\ &= \frac{x_n}{2} + \frac{a}{2x_n} \end{aligned}$$

proof of superlinear conv. of secant method.

Similarly to Newton's method, subtract the zero ξ from the method

$$x_{n+1}-\xi = x_n-\xi - f(x_n)\,\frac{x_n-x_{n-1}}{f(x_n)-f(x_{n-1})} \qquad (*)$$

Use MVT : $f(x_n)-f(x_{n-1}) = f'(\eta_n)(x_n-x_{n-1})$

Plug this in, we obtain

$$e_{n+1} = e_n - f(x_n)/f'(\eta_n)$$

$$f(x_n) = \underbrace{f(\xi)}_{=0} + f'(\xi)(\xi - x_n) + f''(\alpha_n)/2 \cdot (\xi - x_n)^2$$

$$= e_n \Big(\underbrace{1 - \underbrace{\frac{f'(\xi)}{f'(\eta_n)}}_{\to 1} + \underbrace{\frac{f''(\alpha_n)}{2}(\xi - x_n)}_{\to 0}}_{\to 0} \Big).$$

We can still extract something tending to 0!

This time we continue $(*)$

$$e_{n+1} = \frac{(f(x_n) - f(x_{n-1}))e_n - f(x_n)\overbrace{(x_n - x_{n-1})}^{\pm\xi \to e_n - e_{n-1}}}{f(x_n) - f(x_{n-1})}$$

$$= \frac{f(x_n)e_{n-1} - f(x_{n-1})e_n}{f(x_n) - f(x_{n-1})}$$

$$= e_n e_{n-1} \frac{\overbrace{\frac{f(x_n)}{e_n} - \frac{f(x_{n-1})}{e_{n-1}}}^{\to \text{“} \simeq f'(x_n) - f'(x_{n-1}) \text{”}}}{f(x_n) - f(x_{n-1})}$$

(see below)

$$= e_n e_{n-1} \underbrace{\frac{(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})}}_{\to C \text{ as } e_n \to 0} \frac{f''(\xi) \cdot (O(e_n^2) + O(e_{n-1}^2))}{}$$

$$= e_n e_{n-1} \frac{f''(\xi)}{f'(\beta_n)} \cdot (O(e_n^2) + (e_{n-1}^2))$$

$$f(x_n) = f(\xi) + f'(\xi)(\xi - x_n) + \frac{f''(\xi)}{2}(\xi - x_n)^2 + \cdots$$

$$\Rightarrow f(x_n)/e_n = -f'(\xi) - \frac{f''(\xi)}{2}(\xi - x_n) + O(e_n^2)$$

$$f(x_{n-1}) = f(\xi) + f'(\xi)(\xi - x_{n-1}) + \frac{f''(\xi)}{2}(\xi - x_{n-1})^2 + O(e_{n-1}^2)$$

$$\Rightarrow f(x_{n-1})/e_{n-1} = -f'(\xi) - \frac{f''(\xi)}{2}(\xi - x_{n-1}) + \cdots$$

$$\Rightarrow f(x_n)/e_n - f(x_{n-1})/e_{n-1} = f''(\xi)(x_n - x_{n-1}) + O(e_n^2) + O(e_{n-1}^2)$$

In short,

$$e_{n+1} \approx C\, e_n e_{n-1} \qquad \text{———} (**)$$

Let's assume that a fixed rate of conver is available and call it φ:

$$|e_{n+1}| \approx A|e_n|^{\varphi} \qquad \text{when } |e_n| \text{ is small}$$

Plug this into $(**)$.

$$\begin{array}{ccc} |e_{n+1}| & \approx & C|e_n||e_{n-1}| \\ \| & & \| \\ A|e_n|^{\varphi} & & CA|e_{n-1}|^{\varphi+1} \\ \| & & \\ A(A|e_{n-1}|^{\varphi})^{\varphi} & & \\ \| & & \\ A^{1+\varphi}|e_{n-1}|^{\varphi^2} & & \end{array}$$

$$\Rightarrow \quad \underbrace{|e_{n-1}|}_{\text{tending to 0 or } \infty \text{ unless } \varphi^2-\varphi-1=0}{}^{\varphi^2-\varphi-1} \approx CA^{-\varphi} \quad (\text{constant})$$

$$\Rightarrow \quad \varphi = \frac{1 \pm \sqrt{5}}{2}$$

$$\Rightarrow \quad \varphi = \frac{1+\sqrt{5}}{2} \qquad (\text{take positive one})$$

A common trick using big oh.

$\boxed{\text{Def}}$ If there is a neighborhood I of a and $C>0$ such that

$$|f(x)| \leq C(g(x)) \quad \forall x \in I \setminus \{a\}$$

we denote $f(x) = O(g(x))$ as $x \to a$ and say "$f(x)$ is big oh of $g(x)$" or "$f(x)$ is of order of $g(x)$".

Intuition: $f(x)$ is almost constant multiple of $g(x)$ near $x=a$.

Advantage:

give minimal attention to what does not matter much.

- sometimes, big o term does not matter (our current situation)
- other times, everything other than big o term does not matter.

example : part of the proof above

Compare the two scenarios.

1) Using big o (see above)

2) Full details without big o.

$$= e_n e_{n-1} \frac{\overbrace{\frac{f(x_n)}{e_n} - \frac{f(x_{n-1})}{e_{n-1}}}^{\to\ \approx\ \text{"}f'(x_n) - f'(x_{n-1})\text{"}\ \text{(see below)}}}{f(x_n)-f(x_{n-1})}$$

$$= e_n e_{n-1} \underbrace{\frac{(x_n - x_{n-1})}{f(x_n)-f(x_{n-1})} \frac{f''(\xi) \cdot \left(\frac{f'''(a_n)}{6} e_n^2 - \frac{f'''(a_{n-1})}{6} e_{n-1}^2\right)}{}}_{\to\ C \text{ as } e_n \to 0}$$

$$= e_n e_{n-1} \frac{f''(\xi)}{f'(\beta_n)} \cdot \left(\frac{f'''(a_n)}{6} e_n^2 - \frac{f'''(a_{n-1})}{6} e_{n-1}^2\right)$$

$$f(x_n) = f(\xi) + f'(\xi)(\xi - x_n) + \frac{f''(\xi)}{2}(\xi - x_n)^2 + \frac{f'''(a_n)}{6}(\xi - x_n)^3 \quad a_n \in (x_n, \xi)$$

$$\Rightarrow f(x_n)/e_n = -f'(\xi) - \frac{f''(\xi)}{2}(\xi - x_n) + \frac{f'''(a_n)}{6} e_n^2 \quad a_{n-1} \in (x_{n-1}, \xi)$$

$$f(x_{n-1}) = f(\xi) + f'(\xi)(\xi - x_{n-1}) + \frac{f''(\xi)}{2}(\xi - x_{n-1})^2 + \frac{f'''(a_{n-1})}{6}(\xi - x_n)^3$$

$$\Rightarrow f(x_{n-1})/e_{n-1} = -f'(\xi) - \frac{f''(\xi)}{2}(\xi - x_{n-1}) + \frac{f'''(a_{n-1})}{6} e_{n-1}^2$$

$$\Rightarrow f(x_n)/e_n - f(x_{n-1})/e_{n-1} = f''(\xi)(x_n - x_{n-1}) + \frac{f'''(a_n)}{6} e_n^2 - \frac{f'''(a_{n-1})}{6} e_{n-1}^2$$

(proof of contraction mapping thm)

$$\begin{aligned}|x_{n+1}-x_n| &= |F(x_n)-F(x_{n-1})| \\ &\le \lambda\,|x_n-x_{n-1}| \\ &\le \lambda^2\,|x_{n-1}-x_{n-2}| \\ &\vdots \\ &\le \lambda^n\,|x_1-x_0|\end{aligned}$$

Observe $x_n = S_n - S_{n-1}$ $(n \ge 1)$

$$= (x_n - x_{n-1}) + (x_{n-1} - x_{n-2}) + \cdots + (x_1 - x_0) + x_0,$$

where $S_n = \sum_{j=0}^{n} x_j$.

$$\sum_{n=0}^{\infty} |x_{n+1}-x_n| \le \sum_{n=0}^{\infty} \lambda^n |x_1 - x_0|$$

$$= |x_1 - x_0| \frac{1}{1-\lambda} \quad \text{(convergent)}$$

$\Rightarrow \sum_{n=0}^{\infty} (x_{n+1}-x_n)$ converges.

(absolute conv. implies conv.)

$$\Rightarrow \lim_{n\to\infty} x_n = \lim_{n\to\infty}\left(\sum_{n=0}^{\infty}(x_{n+1}-x_n)+x_0\right) = x$$

Then, take $n \to \infty$ in $x_{n+1} = F(x_n)$:

$$x = F(x).$$

/* Contraction is continuous

If $|x-y| < \varepsilon/\lambda$, then

$$|F(x) - F(y)| \le \lambda |x-y| < \lambda \cdot \frac{\varepsilon}{\lambda} = \varepsilon. \text{ */}$$

Therefore, the limit x is the fixed point of F.

- $x \in D$ b/c D is closed.
- Fixed point is unique since if y is another one,

$$|x-y| = |F(x) - F(y)| \le \lambda |x-y|$$

$$\Rightarrow (1-\lambda)|x-y| \le 0 \Rightarrow x = y.$$

example of contraction mapping

diverging $\lambda=1$.