Finding zeros of $f: \mathbb{R} \to \mathbb{R}$ or $f: \mathbb{R}^n \to \mathbb{R}$

- $f: \mathbb{R} \to \mathbb{R}$ — zeros of polynomials there are known methods
  - if f is not a polynomial then finding a zero(s) becomes non-trivial

- Optimization :- Any minimization / maximization problem reduces to understanding critical points. More precisely

$F: \mathbb{R}^n \to \mathbb{R}$ $\quad \max_{x \in A} F(x)$ or $\min_{x \in B} F(x)$

One needs to find

$$\{x : \nabla F(x) = 0\}$$

we will explore

$f: \mathbb{R} \to \mathbb{R}$	$F: \mathbb{R}^n \to \mathbb{R}$; $\min_{x \in A} F(x)$
• Bisection method	• Golden Section Search
- Newton-Raphson method	• Generalization of Newton-Raphson
• Secant method	• Gradient descent
– Convergence	• Hill climbing
– rate	• Simulated Annealing
– limitations	- Methods, usefulness & implementation on R
– Both theoretical aspects & implementation on R	

Bisection Method:-

$f: [a,b] \to \mathbb{R}$

Assume $f(a) < 0$, $f(b) > 0$, f is continous.

Aim:- find c : $f(c) = 0$, $c \in (a,b)$

Algorithm:

$a = ..$

$b = ..$

for k in $1:n$

$x_k = \frac{a+b}{2}$ [bisect the interval]

if $\text{sign} f(x_k) = \text{sign} f(a)$

$a = x_m$

if $\text{sign} f(x_k) = \text{sign} f(b)$

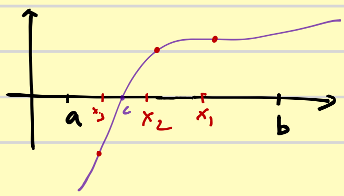
$b = x_m$

end

return (x_n).

One can decide on n by the following

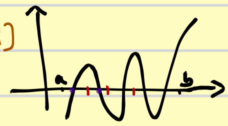
- stop at some arbitrary large n
- stop when $|f(x_k)| < tol$



let $c \in (a,b)$ $f(c)=0$. observe

$$|x_k - c| \le |a_k - b_k| = \frac{1}{2^k}|a-b|$$

- Each bisection step - discovers a new correct digit in binary expansion of c
- Clearly $x_k \to c$ as $k \to \infty$; the method will converge under hypothesis
- [no control]



- could converge to one of the roots

- [NOT really fast] only guarantees that

$$|x_{k+1} - c| < \frac{1}{2} |x_k - c|$$

$\boxed{\text{fixed}}$

[linear Convergence] will explain below

Extensions:- $f: \mathbb{R} \to \mathbb{R}$ continuous; one can use bracketing to designate an interval to apply bisection.

Algorithm:- f x_{min}, x_{max}, n

$d = \frac{x_{max} - x_{min}}{n}$

$a = x_{min}$

while $i < n$

  $i \to i+1$

  $b = a + d$

  if f changes sign in $[a,b]$, save

  end

  $a = b$

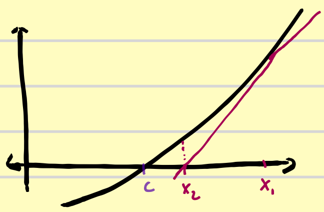
end

Newton-Raphson

- important method and it is the basis of most optimization solutions.

Assume: $f: \mathbb{R} \to \mathbb{R}$ and f is twice differentiable

Idea :-



- Drop tangent line at $(x_n, f(x_n))$
- x-intercept of tangent line $= x_{n+1}$

Algorithm :-

$y - f(x_1) = (x - x_1) f'(x_1)$

$y=0$ $\quad x_2 = -\dfrac{f(x_1)}{f'(x_1)} + x_1$

x_1 = initialize
for k in 2:n
$\quad x_k = x_{k-1} - \dfrac{f(x_{k-1})}{f'(x_{k-1})}$
$\quad$ at convergence criteria stop
end
return (x_k)

Example (already seen): $f(x) = x^2 - a \qquad a > 0$

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

Newton's Method reduces to: $x_1 > \sqrt{a}$

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

$$\Rightarrow \quad x_{k+1} = x_k - \frac{x_k^2 - a}{2x_k}$$

$$\Rightarrow \quad x_{k+1} = \frac{1}{2}\left(x_k - \frac{a}{x_k} \right)$$

Already seen the following:

- $x_k \to \sqrt{a}$ as $k \to \infty$

  & $x_{k+1} - \sqrt{a} \le \dfrac{(x_k - \sqrt{a})^2}{2\sqrt{a}} \quad - \forall k \geq 1$

  [compare with Bisection
  - Quadratic Convergence]
  - will explain later

- In the case $a = 3$, $x_1 = 2$ we saw

  $$x_6 - \sqrt{a} < 10^{-31}$$

  Q:- Is this characteristic of the method?

- Assume:
  - $f : \mathbb{R} \to \mathbb{R}$; f - twice differentiable
  - $f(c) = 0$
  - $\exists \delta > 0$: $f'(x) \neq 0 \quad \forall x \in (c - \delta, c + \delta)$
  - f'' - continuous.

If $x_k \in (c-\delta, c+\delta) \quad \forall k \geq 1$

we also have $x_{k+1} = x_k - \dfrac{f(x_k)}{f'(x_k)}$ — ②

need $f'(x_k) \neq 0$

By Taylor's theorem $\exists\ \xi_k$ between x_k and c

such that

$$f(c) = f(x_k) + (c - x_k) f'(x_k) + \frac{1}{2}(c - x_k)^2 f''(\xi_k) \quad \text{— ①}$$

$$f(c) = 0 \quad \text{— ③}$$

Substitute:

③ in ① & value of $f(x_k)$ from ② into ①

to get

$$0 = (x_{k+1} - x_k) f'(x_k) + (c - x_k) f'(x_k) + \frac{1}{2}(c - x_k)^2 f''(\xi_k)$$

$$\Rightarrow \quad (x_{k+1} - c) = \frac{1}{2} \frac{f''(\xi_k)}{f'(x_k)} (x_k - c)^2 \quad \text{— ④}$$

$$\Rightarrow \quad (x_{k+1} - c) \leq M_1 (x_k - c)^2$$

$$\Rightarrow \quad (x_{k+1} - c) \leq \frac{[M_1 (x_0 - c)]^{2^k}}{M_1}$$

$\therefore$

If $|M_1 (x_0 - c)| < 1$ then $x_k \to c$ as $k \to \infty$.

From the above calculation we can postulate

Theorem (a Sufficient Condition) :-

$f : \mathbb{R} \to \mathbb{R}$; f - twice differentiable

- $f(c) = 0$
- $\exists\, 0 < \delta$, $f'(x) \neq 0 \quad \forall x \in (c-\delta, c+\delta)$
- f'' - continuous in $(c-\delta, c+\delta)$

$\exists\ \eta > 0$ s.t. if $x_0 \in (c-\eta, c+\eta)$

ⓐ $x_{k+1} = x_k - \dfrac{f(x_k)}{f'(x_k)}$ [well defined $\forall k \geq 1$]

ⓑ $(x_{k+1} - c) < M\,(x_k - c)^2$

& $x_{k+1} - c \leq \dfrac{(M\,(x_0 - c))^{2^k}}{M}$

for any $M > \frac{1}{2}\dfrac{|f''(c)|}{|f'(c)|}$

In particular $x_k \longrightarrow c$ as $k \to \infty$
(quadratically).

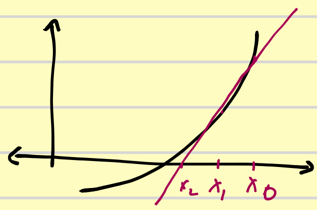
Secant Method :.

$f : \mathbb{R} \to \mathbb{R}$, Suppose Computing the derivatives are hard. We discussed last week, finite difference approximation to derivative.

$x, y \in \mathbb{R}$, $x \neq y$ set

$$f[x,y] := \frac{f(x) - f(y)}{x - y}$$

Idea :-



- Drop a secant line from x_0, x_1
- x-intercept of line $= x_2$

$$x_2 = x_1 - \frac{f(x_1)}{f[x_1, x_0]}$$

Algorithm :-

$x_1 =$ initialize

for k in $2:n$

$$x_k = x_{k-1} - \frac{f(x_{k-1})}{f[x_{k-1}, x_{k-2}]}$$

at convergence criteria stop

end

return (x_k)

Assume :
- $f : \mathbb{R} \to \mathbb{R}$; f - twice differentiable
- $f(c) = 0$
- $\exists \delta > 0$: $f'(x) \neq 0$ $\forall x \in (c - \delta, c + \delta)$
- f'' - continuous.

$n \geq 1 \qquad x_{n+1} = x_n - \dfrac{f(x_n)}{f[x_n, x_{n-1}]} \quad — ①$

let $\quad p(x) = f(x_n) + f[x_{n-1}, x_n](x - x_n)$

$f(x) - p(x)$

$= f(x) - \left[f(x_n) + \dfrac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} (x - x_n) \right]$

$= \left[\dfrac{f(x) - f(x_n)}{x - x_n} - \dfrac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} \right] \cdot (x - x_n)$

$= f[x; x_n, x_{n-1}] (x - x_n)(x - x_{n-1})$

claim ⊛ $\quad \dfrac{f''(\xi)}{2} (x - x_n)(x - x_{n-1})$

where $\quad \xi \in [a,b] \supseteq \{x, x_n, x_{n-1}\}$.

Applying above at $x = c$

$0 = f(x_n) + f[x_{n-1}, x_n](c - x_n) + \dfrac{f''(\xi)}{2} (c - x_n)(c - x_{n-1}) \quad — ②$

① yields

$0 = f(x_n) + f[x_{n-1}, x_n](x_{n+1} - x_n) \quad — ③$.

② - ③ $\Rightarrow$

$c - x_{n+1} = \dfrac{1}{2} \dfrac{f''(\xi)}{f[x_{n-1}, x_n]} (c - x_n)(c - x_{n-1})$