Newton's method for root-finding

Spoiler: Quadratic convergence when close to a root. (faster than bisection)
Issue: not guaranteed always to converge

- Want to find a zero of some function f
- Let r be a zero of $f \iff f(r) = 0$
- If the Taylor series makes sense, i.e., f'' exists and if our current guess for the zero is $x = r - h$
$\iff r = x + h$

$$\underset{f(r)=0}{f(r)} = f(x+h) \underset{\text{Taylor}}{=} \underbrace{f(x) + hf'(x)}_{\text{linear approximation}} + \underbrace{\frac{h^2 f''(\xi)}{2}}_{O(h^2)}$$

So: $0 = f(x) + hf'(x) + O(h^2)$

For small h, we approximate. $0 \approx f(x) + hf'(x)$

So $h \approx \dfrac{-f(x)}{f'(x)}$ (Remember, we know x so if we know h, we know $r = h + x$!)

and $r \approx x - f(x)/f'(x)$

Now, our new approximation is $x \leftarrow x - f(x)/f'(x)$

That is we have the iteration

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

Simple Pseudo-code

```
INPUT x, M
 y ← f(x)

FOR k=1 to M DO
          x ← x - y/f'(x)
          y ← f(x)
END DO
```

Pseudocode with stopping criteria

INPUT $x_0, M, \delta, \varepsilon$

$v \leftarrow f(x_0)$

IF $|v| < \varepsilon$ THEN STOP

FOR $k = 1$ to M DO

$x_1 \leftarrow x_0 - y / f'(x_0)$

$v \leftarrow f(x_1)$

IF $|x_1 - x_0| < \delta$ or $|v| < \varepsilon$ THEN STOP

$x_0 \leftarrow x_1$

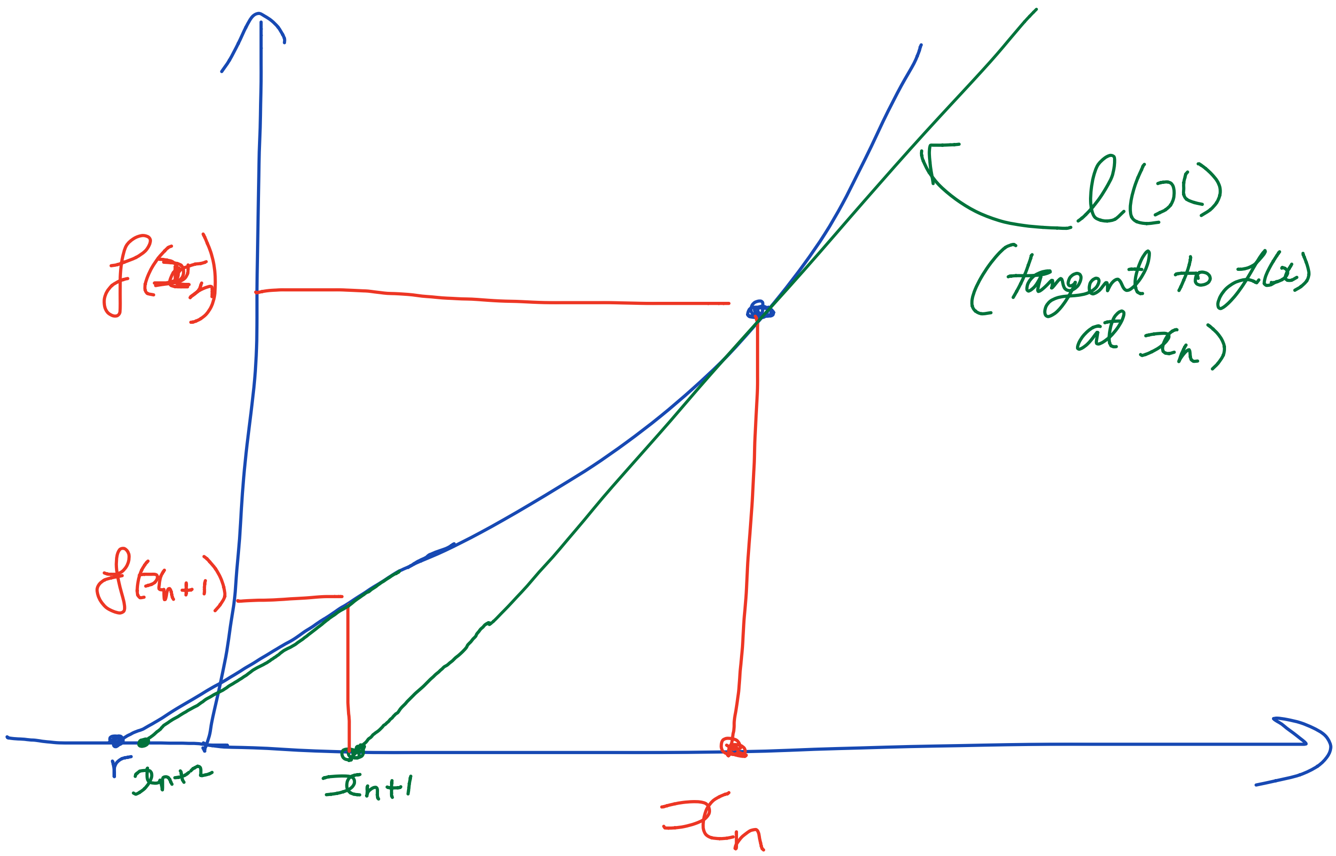
END DO

Interpretation:

$$f(x) = f(x_n) + f'(x_n)\cdot(x-x_n) + \cdots\cdots$$

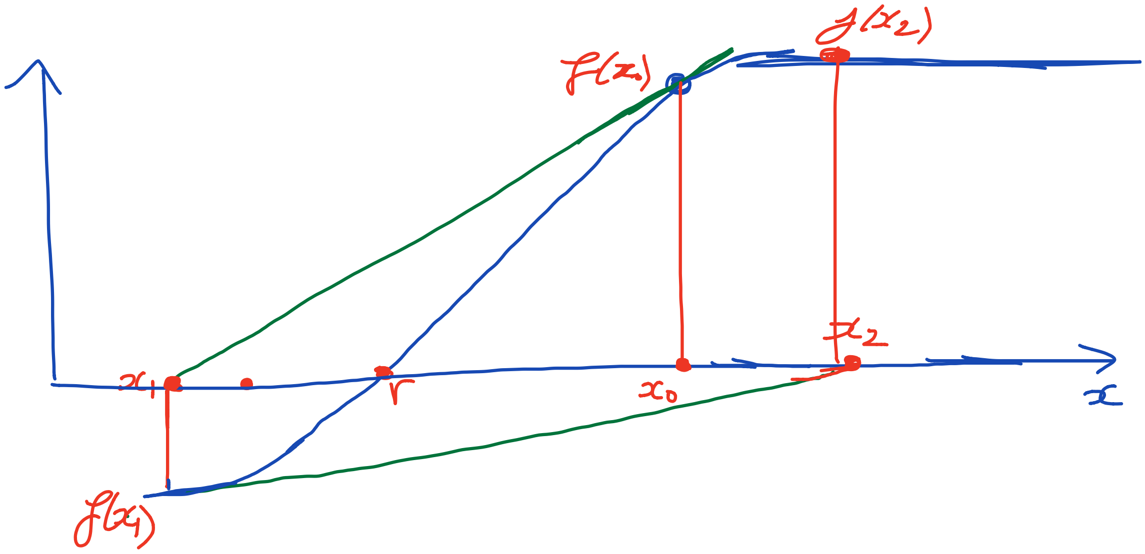
$$\ell(x) = f(x) + f'(x_n)(x-x_n)$$

x_{n+1} is a root of $\ell(x)$.



Issue: If x_0 is not very close to a zero, or if the graph of f is "not nice", Newton's method may fail.

example:



Error Analysis:

Define the error at n'th step: $e_n = x_n - r$

Assume $f \in C^2(\mathbb{R})$ & $f(r)=0$ but $f'(r) \neq 0$ (simple zero)

then: $\quad e_{n+1} = x_{n+1} - r$

$$= x_n - \frac{f(x_n)}{f'(x_n)} - r \quad \text{(by def'n of Newton's method)}$$

$$= \underbrace{x_n - r}_{e_n} - f(x_n)/f'(x_n)$$

$$= e_n - f(x_n)/f'(x_n) = \frac{e_n f'(x_n) - f(x_n)}{f'(x_n)} \quad \text{——— ①}$$

But $f(r) = 0 = f(x_n - e_n) = f(x_n) - e_n f'(x_n) + \frac{e_n^2 f''(\xi_n)}{2}$

for some ξ_n between x_n & r $\Rightarrow e_n f'(x_n) - f(x_n) = \frac{e_n^2 f''(\xi_n)}{2}$

——— ②

① & ② $\Rightarrow$ $e_{n+1} = e_n^2 \dfrac{f''(\xi_n)}{f'(x_n)\cdot 2} \approx C e_n^2$ when x_n is close to r.

③

quadratic convergence

So we need to understand when $f''(\xi_n)/f'(x_n)$ is small

Let $c(\delta) = \frac{1}{2} \max_{|x-r|<\delta} |f''(x)| \Big/ \min_{|x-r|<\delta} |f'(x)|$ for $\delta > 0$

Pick δ small enough so $\min_{|x-r|<\delta} |f'(x)| > 0$ & $\delta c(\delta) < 1$

Let $\rho = \delta c(\delta)$ & suppose $|x_0 - r| \leqslant \delta$

$\Leftrightarrow |e_0| \leqslant \delta \Rightarrow |\xi_0 - r| \leqslant \delta \Rightarrow \frac{1}{2} |f''(\xi_0)/f'(x_0)| \leqslant c(\delta)$

③ $\Rightarrow |e_1| \leqslant e_0^2 c(\delta) = |e_0| \underbrace{|e_0| c(\delta)}_{\leqslant \delta c(\delta)} \leqslant \rho |e_0| < |e_0| \leqslant \delta$

Since $|e_1| = |x_1 - r| \leqslant \delta$

we can do the same thing for e_2

$\Rightarrow |e_2| \leqslant \rho |e_1| \leqslant \rho^2 |e_0|$

$\vdots$

$|e_n| \leqslant \rho^n |e_0| \xrightarrow[n \to \infty]{} 0$ (bec $\rho < 1$

Theorem: Suppose $f \in C^2(\mathbb{R})$, $f(r)=0$, $f'(r) \neq 0$. Then $\exists\, \delta > 0$ and a $\rho < 1$, such that if $|x_0 - r| < \delta$ Newton's method started at x_0 yields

$$|x_{n+1} - r| \leq \rho (x_n - r)^2 \quad \text{for } n \geq 0.$$

Theorem: If $f \in C^2(\mathbb{R})$, increasing, convex ($f'' > 0$), has a zero then the zero is unique & the Newton iteration will converge to it from any starting point.

e.g. $f(x) = x^2 - R$ has a root at $x = \sqrt{R}$ so we can use Newton's method to find it

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - R}{2x_n}$$

$$= \frac{1}{2}\left(x_n + \frac{R}{x_n}\right)$$

~~Systems of nonlinear equations~~

(will return to this later, you can skip for now)

Recall that Newton's iteration was derived from linearizing the $f^{\underline{n}}$ (i.e. using 1st order Taylor approx)

Same idea for vector-valued functions of many variables.

$$\text{want to solve: } \left.\begin{cases} f_1(x_1, \dots, x_n) = 0 \\ f_2(x_1, x_2, \dots, x_n) = 0 \\ \quad \vdots \\ f_n(x_1, x_2, \dots, x_n) = 0 \end{cases}\right\} \begin{array}{l} n\text{-equations} \\ \to n \\ \text{unknowns} \end{array}$$

$$\iff F(X) = 0 \quad \text{where } X = (x_1, x_2, \dots, x_n)^T$$
$$F = (f_1, f_2, \dots, f_n)^T$$

Starting at an estimate X:

$$F(X+H) \approx F(X) + \underbrace{F'(X)}_{\substack{n\times n \\ \text{Jacobian}}} H$$

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

If $X+H$ is a root then $F(X+H) = \underbrace{0 \approx F(X) + F'(X)H}_{\substack{\text{so, we prefer to solve the} \\ \text{system by, e.g.,} \\ \text{Gaussian elimination}}}$

So $H = -\underbrace{(F'(X))^{-1}}_{\substack{\text{can be expensive} \\ \text{to invert large matrices}}} F(X)$

$$\boxed{\text{Newton's method: } X^{(n+1)} = X^{(n)} + H^{(n)} \text{ where } H^{(n)} \text{ satisfies } F'(X^{(n)})\, H^{(n)} = -F(X^{(n)})}$$

Equivalently: $X^{(n+1)} = X^{(n)} - \left(F'(X^{(n)})\right)^{-1} F(X^{(n)})$

Example: To solve: $\begin{cases} xy = z^2+1 \\ xyz + y^2 = x^2+2 \\ e^x + z = e^y + 3 \end{cases}$

we set-up $F(X) = \begin{pmatrix} f_1(x,y,z) \\ f_2(x,y,z) \\ f_3(x,y,z) \end{pmatrix} = \begin{pmatrix} xy - z^2 - 1 \\ xyz + y^2 - x^2 - 2 \\ e^x + z - e^y - 3 \end{pmatrix}$

$$\Rightarrow F'(X) = \begin{pmatrix} y & x & -2z \\ yz-2x & xz+2y & xy \\ e^x & -e^y & 1 \end{pmatrix}$$

$\vdots$

Secant Method

Newton: $x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)}$

Secant: $x_{n+1} = x_n - \dfrac{f(x_n)}{\dfrac{f(x_n)-f(x_{n-1})}{x_n - x_{n-1}}}$ → approx to $f'(x_n)$

$\iff \quad x_{n+1} = x_n - \dfrac{f(x_n)\,(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})}$

$n \geq 1$

Remark: Need two initial points

Remark: $|e_{n+1}| \approx A\,|e_n|^{(1+\sqrt{5})/2}$

A ↑ const. >0

$A = \left|\dfrac{f''(r)}{2f'(r)}\right|^{\frac{\sqrt{5}-1}{2}}$