Section 3.2: Newton's method for root-finding (Newton-Raphson iteration)

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 $x = r - h$

$$\iff r = x + h$$

$$f(r) = f(x+h) = f(x) + hf'(x) + \frac{h^2 f''(y)}{2}$$

$f(r) = 0$; Taylor ; linear approximation ; $O(h^2)$

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

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

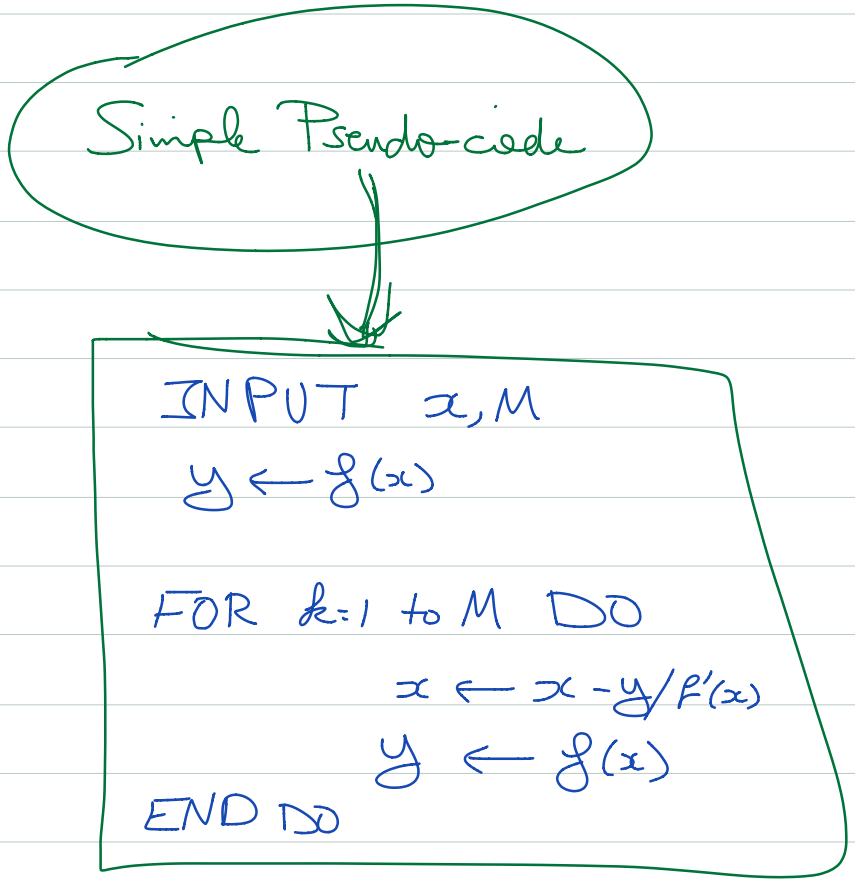
So $h \approx \frac{-f(x)}{f'(x)}$ (Remember, we know x so if we know h, we know $r = h + x$!)

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

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



Pseudocode with stopping criteria

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

$y \leftarrow f(x_0)$

IF $|y| < \varepsilon$ THEN STOP

FOR $k = 1$ to M DO

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

$y \leftarrow f(x_1)$

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

$x_0 \leftarrow x_1$

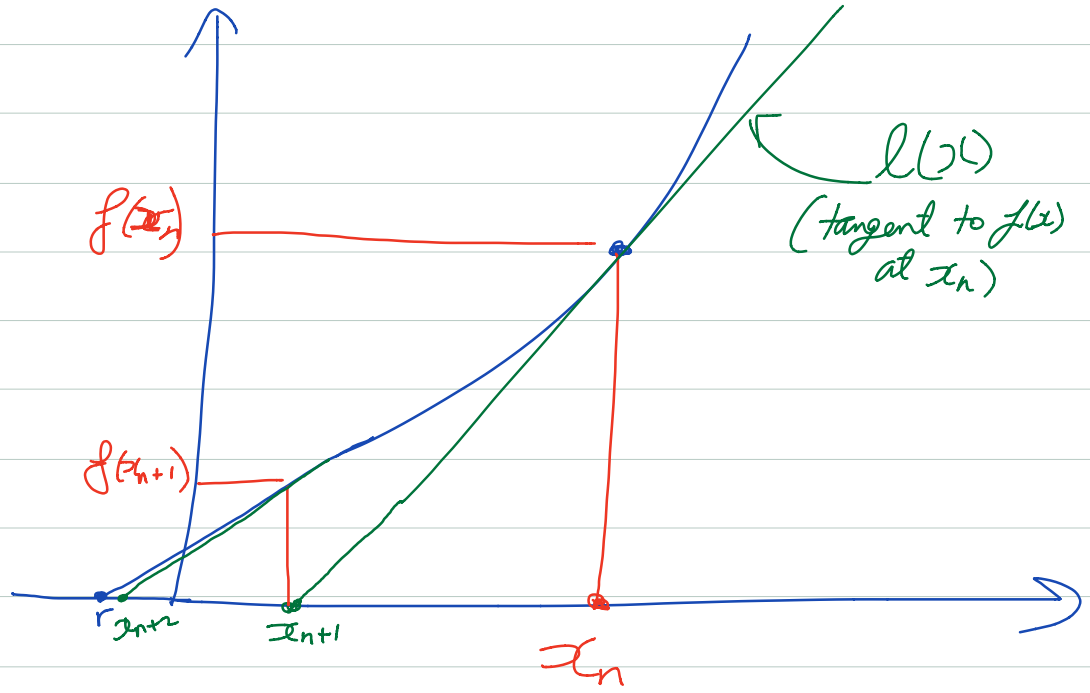
END DO

Interpretation:

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

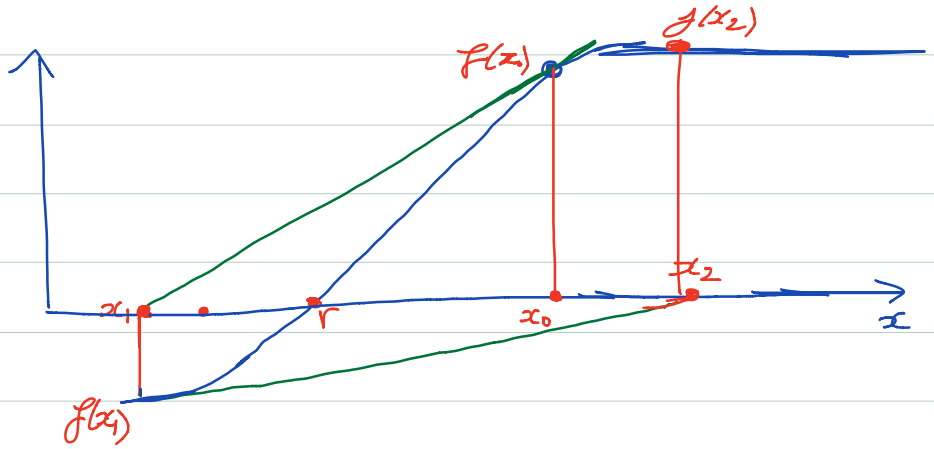
$$\ell(x) = f(x) + f'(x_n)(x - x_n)$$

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

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

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

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

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

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

——— (2)

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

③

$C e_n^2$: quadratic convergence

So we need to understand when $f''(\varphi_n)/f'(x_n)$ 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 $|x_0 - r| \leq \delta$

$\Leftrightarrow |e_0| \leq \delta \Rightarrow |\varphi_0 - r| \leq \delta \Rightarrow \frac{1}{2} |f''(\varphi_0)/f'(x_0)| \leq c(\delta)$

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

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

we can do the same thing for e_2

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

$\vdots$

$|e_n| \leq \rho^n |e_0| \xrightarrow[n\to\infty]{} 0$ (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:

Recall that Newton's iteration was derived from linearizing the f.t (i.e. using 1st ordered Taylor approx)

Same idea for vector-valued functions of many variables.

want to solve: $\begin{cases} f_1(x_1, \dots, x_n) = 0 \\ f_2(x_1, x_2 \dots, x_n) = 0 \\ \vdots \\ f_n(x_1, x_2 \dots, x_n) = 0 \end{cases}$ n equations, n unknowns

$\iff F(X) = 0$ where $X = (x_1, x_2, \dots, x_n)^T$

$F = (f_1, f_2, \dots, f_n)^T$

Starting at an estimate X:

$F(X+H) \approx F(X) + F'(X)H$

$F'(X)$: $n \times n$ Jacobian

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

If $X+H$ is a root then $F(X+H) = 0 \approx F(X) + F'(X)H$

So $H = -(F'(X))^{-1} F(X)$

can be expensive to invert large matrices

So, we prefer to solve the system by, e.g., Gaussian elimination

Newton's method: $X^{(n+1)} = X^{(n)} + H^{(n)}$

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

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

$\vdots$

HW: Sec. 3.2: 5, 9, 13, 22, 23

Sec 3.3: 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 $f'(x_n)$

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

$n \geq 1$

Remark: Need two initial points

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

↑ const. > 0

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

Section 3.4 Fixed points & Functional iteration

Newton's method: $x_{n+1} = x_n - \underbrace{\frac{f(x_n)}{f'(x_n)}}_{\text{call this } F(x_n)}$

So, Newton's method is a functional iteration

$$x_{n+1} = F(x_n)$$

Suppose (in general now) that
$\lim_{n\to\infty} x_n$ exists & $\lim_{n\to\infty} x_n = s$

Suppose also that F is continuous

Then $\lim_{n\to\infty} F(x_n) = F\left(\lim_{n\to\infty} x_n\right) = F(s)$

$\| \left(\text{since } x_{n+1} = F(x_n)\right)$

$\lim_{n\to\infty} x_{n+1} = s$

So $s = F(s)$!!

Such a point is called a fixed pt. of F.

Very important & interesting problems (e.g. in optimization, algorithm design, diff. equations ....) can be reduced to finding the fixed pts of a function F.

Simple (but important case)

$F: \quad C \longrightarrow C$ where C is a closed subset of $\mathbb{R}$.

Contractive mapping:

$$|F(x) - F(y)| \leq \lambda |x - y|$$

for some $\lambda < 1$

(can you see why it's called contractive?)

Theorem: Let $C \subseteq \mathbb{R}$ be closed. If $F: C \to C$ is contractive, then F has a unique fixed pt S. Moreover

$$S = \lim_{n \to \infty} x_{n+1} \text{ where } x_{n+1} = F(x_n)$$

for any starting pt $x_0 \in C$.

Proof: want to show that $(x_n)_{n=0}^{\infty}$ converges

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

$$= \left(\sum_{i=1}^{n} (x_i - x_{i-1}) \right) + x_0$$

want this sequence to converge as $n \to \infty$

It suffices to show that $\sum_{i=1}^{n} |x_i - x_{i-1}|$ converges

but

$$|x_{i+1} - x_i| = |F(x_i) - F(x_{i-1})|$$

contractive map $\Rightarrow \leq \lambda |x_i - x_{i-1}|$

$$\vdots \leq \lambda^i |x_1 - x_0|$$

So $\sum_{i=1}^{\infty} |x_i - x_{i-1}| \leq \sum_{i=1}^{\infty} \lambda^i |x_1 - x_0| = \frac{1}{1-\lambda} |x_1 - x_0|$

$\Rightarrow$ the sequence converges, so let $s = \lim_{n \to \infty} x_n$ and note that $s = F(s)$.

To see that s is unique, suppose t is a different fixed pt. so $t = F(t)$

$$|t - s| = |F(t) - F(s)| \leq \lambda |t - s|$$

but $\lambda < 1$ so $t = s$. ∎

Exercise: Let $F(x) = a + b\sin(x)$ for some $a \in \mathbb{R}$ & $b \in \mathbb{R}$. For what values of a & b is F contractive? In that case, write an iteration to find the fixed pt of F.

Order of Convergence: Suppose $\begin{cases} x_{n+1} = F(x_n) \\ F(s) = s \end{cases}$

Let $e_n = x_n - s$

(then $\lim_{n\to\infty} e_n = 0$) & the order of convergence is the smallest integer k s.t. $F^{(k)}(s) \neq 0$.