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

Spoiler: Quadratic convergence when close to a root. (Faster than linear)
Issue: Not guaranteed always to converge

- Want to find a zero $x_0$ of some function $f$

- Let $r$ be a zero of $g$ $\iff g(r) = 0$

- If the Taylor series makes sense, i.e., $f''$ exists
  and if our current guess for the zero is $x = r - h$

  $r = x + h$

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

  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 $h \approx f(x)$ (Remember, we know $a$

  so if we know $h$, we know $r = h + x$)

  and $r \approx x - \frac{f(x)}{f'(x)}$
Now, our new approximation is \( x \leftarrow x - \frac{g(x)}{g'(x)} \)

That is, we have the iteration

\[
 x_{n+1} = x_n - \frac{g(x_n)}{g'(x_n)}
\]

**Simple Pseudo-code**

```plaintext
INPUT x, M
y <- g(x)

FOR k=1 to M DO
    x <- x - y / f'(x)
    y <- g(x)
END DO
```
Pseudocode with stopping criteria

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

$\nu \leftarrow g(x_0)$

IF $|\nu| < \varepsilon$ THEN STOP

FOR $k = 1 \text{ to } M$ DO

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

$\nu \leftarrow g(x_1)$

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

$x_0 \leftarrow x_1$

END DO
Interpretation:

\[ f(x) = f(x_n) + f'(x_n)(x-x_n) + \ldots \]

\[ l(x) = f(x) + f'(x_n)(x-x_n) \]

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

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 design of Newton's method)

$= x_n - r - \frac{f(x_n)}{f'(x_n)}$

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

$= e_n - \frac{f'(x_n)}{f'(x_n)} = \frac{e_n f'(x_n) - f(x_n)}{f'(x_n)}$ \hspace{1cm} (1)

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

For some $\xi_n$ between $x_n$ & $r$ \Rightarrow $e_n f'(x_n) - f(x_n) = \frac{e_n^2 f''(\xi_n)}{2}$ \hspace{1cm} (2)
1 \& 2 \Rightarrow e_{n+1} = e_{n} \frac{f''(r_{n})}{f'(r_{n})^2} \approx C e_{n} \text{ when } x_{n} \text{ is close to } r_{n}.

\text{quadratic convergence}

So we need to understand when \( \frac{f''(r_{n})}{f'(r_{n})} \) is small.

Let \( C(s) = \frac{1}{2} \max_{|x-r| < s} \frac{|f''(x)|}{\min_{|x-r| < s} |f'(x)|} \) for \( s > 0 \).

Pick \( s \) small enough so \( \min_{|x-r| < s} |f'(x)| > 0 \) & \( s C(s) < 1 \).

Let \( p = s C(s) \) & suppose \( |x_{0} - r| < s \).

\( \Rightarrow |e_{0}| < s \Rightarrow |r_{0} - r| < s \Rightarrow \frac{1}{2} \left| \frac{f''(r_{0})}{f'(r_{0})} \right| < C(s) \).

\( \Rightarrow |e_{1}| \leq s^{2} C(s) = |e_{0}| |e_{0}| C(s) \leq p |e_{0}| < |e_{0}| \leq s \leq s C(s) \).

Since \( |e_{1}| = |x_{1} - r| < s \), we can do the same thing for \( e_{2} \).

\( \Rightarrow |e_{2}| < p |e_{1}| \leq p^{2} |e_{0}| \)

\( \vdots \)

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

\[
|x_{n+1} - r| < p |x_n - r|^2 \quad \text{for } n \geq 0.
\]

**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.

\[f(x) = x^2 - R \] 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 \( f^i \) (i.e. using 1st order Taylor approx.)
Same idea for functions of many variables:

\[ \begin{align*}
F_1(x_1, \ldots, x_n) &= 0 \\
F_2(x_1, x_2, \ldots, x_n) &= 0 \\
&\vdots \\
F_n(x_1, x_2, \ldots, x_n) &= 0
\end{align*} \]

\( \Rightarrow F(X) = 0 \) where \( X = (x_1, x_2, \ldots, x_n)^T \)

\[ F = (F_1, F_2, \ldots, F_n)^T \]

Starting at an estimate \( X \):

\[ F(X + H) \approx F(X) + F'(X)H \]

The Jacobian matrix:

\[ \begin{pmatrix}
\frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_n} \\
\frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} & \cdots & \frac{\partial F_2}{\partial x_n} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial F_n}{\partial x_1} & \frac{\partial F_n}{\partial x_2} & \cdots & \frac{\partial F_n}{\partial x_n}
\end{pmatrix} \]

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

So \( H = -\left(F'(X)\right)^{-1}F(X) \)

This 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: \( X^{(n+1)} = X^{(n)} - (F'(X^{(n)}))^{-1} F(X^{(n)}) \)

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

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

\[ \Rightarrow F'(X) = \left( \begin{array}{ccc} y & x & -2z \\ y^2 - 2x & x^2 + 2y & 2y \\ e^x & -e^z & 1 \end{array} \right) \]
Sec 3.3: Secant Method

Newton: \[ x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \]

Secant: \[ x_{n+1} = x_n - \frac{f(x_n)}{f(x_n) - f(x_{n-1})} \]

\[ x_{n+1} = x_n - \frac{f(x_n)}{F(x_n) - F(x_{n-1})} \]

Rem: Need two initial points

Remark: \[ |e_{n+1}| \approx \frac{A}{\sqrt{2}} |e_n| \]

\[ A = \frac{\frac{8^n(r)}{2f'(r)}}{2^{n+1}} \]

Remark: const. > 0
Section 3.4  Fixed points & Functional iteration

Newton's method: \[ x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \]

So, Newton's method is a functional iteration.

\[ x_{n+1} = F(x_n) \]

Suppose (in general now) that \[ \lim_{n \to \infty} x_n \text{ exists } \] if \[ \lim_{n \to \infty} x_n = s \]

Suppose also that \( F \) is continuous.
Then \( \lim_{n \to \infty} F(x_n) = F(\lim_{n \to \infty} x_n) = F(\alpha) \)

\[ \lim_{n \to \infty} x_{n+1} = \alpha \]

So \( \alpha = F(\alpha) \).

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 pt. of a function \( F \).

Simple (but important case)

\( F : C \to C \) 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

\[
|x_{i+1} - x_i| = |F(x_i) - F(x_{i-1})| \\
\text{contraction map} \Rightarrow \leq \lambda |x_i - x_{i-1}| \\
\Rightarrow |x_i - x_0| \leq \lambda^i |x_1 - x_0|
\]

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

\( \Rightarrow \) 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 an adjacent 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 point \( F \).

Order of Convergence: Suppose \( x_{n+1} = F(x_n) \) \& \( F(s) = s \).

Let \( e_n = x_n - s \).

(then \( \lim_{n \to \infty} e_n = 0 \)) \& the order of convergence is the smallest integer \( k \) \( \geq 1 \) s.t. \( F^{(k)}(s) \neq 0 \).