def: Given a function \( f(x) \), a root is a number \( r \) satisfying \( f(r) = 0 \).

ex: \( f(x) = x^2 - 3 \Rightarrow r = \pm \sqrt{3} \)

question : How can we find the roots of a general function \( f(x) \)?

2.1 bisection method

idea : Find an interval \([a, b]\) such that \( f(a) \) and \( f(b) \) have opposite sign. Then \( f(x) \) has a root in \([a, b]\) (Intermediate Value Theorem, Math 451 - advanced calculus).

Consider the midpoint \( x_0 = \frac{1}{2}(a + b) \). The root \( r \) is contained in either the left subinterval or the right subinterval; to determine which one, compute \( f(x_0) \). Then repeat.

ex: \( f(x) = x^2 - 3 \), \( f(1) = -2 \), \( f(2) = 1 \) ⇒ \( f(x) \) has a root in \([1, 2]\), \( r = 1.73205 \)

| \( n \) | \( a_n \) | \( b_n \) | \( x_n \) | \( f(x_n) \) | \( |r - x_n| \) |
|-------|--------|--------|--------|--------|--------|
| 0     | 1      | 2      | 1.5    | -0.75  | 0.2321 |
| 1     | 1.5    | 2      | 1.75   | 0.0625 | 0.0179 |
| 2     | 1.5    | 1.75   | 1.625  | -0.3594| 0.1071 |
| 3     | 1.625  | 1.75   | 1.6875 | -0.1523| 0.0446 |

bisection method (assume \( f(a) \cdot f(b) < 0 \))

1. \( n = 0 \), \( a_0 = a \), \( b_0 = b \)
2. \( x_n = \frac{1}{2}(a_n + b_n) \) : current estimate of the root
3. if \( f(x_n) \cdot f(a_n) < 0 \), then \( a_{n+1} = a_n \), \( b_{n+1} = x_n \)
4. else \( a_{n+1} = x_n \), \( b_{n+1} = b_n \)
5. set \( n = n + 1 \) and go to line 2
stopping criterion: here are three options

\[ |b_n - a_n| < \epsilon, \quad |f(x_n)| < \epsilon, \quad n = n_{\text{max}} \]

error bound

\[ |r - x_n| \leq \frac{1}{2}|b_n - a_n| = (\frac{1}{2})^2|b_{n-1} - a_{n-1}| = \cdots = (\frac{1}{2})^{(n+1)}|b_0 - a_0| \]

ex: how many steps are needed to ensure that the error is less than \(10^{-3}\)?

\[ [a, b] = [1, 2], \quad |r - x_n| \leq (\frac{1}{2})^{(n+1)}|b_0 - a_0| \leq 10^{-3} \Rightarrow n + 1 \geq 10 \Rightarrow n \geq 9 \]

2.3 fixed-point iteration

Suppose \(f(x) = 0\) is equivalent to \(x = g(x)\). Then \(r\) is a root of \(f(x)\) if and only if \(r\) is a fixed point of \(g(x)\).

![Diagram of fixed-point iteration](image)

We try to solve \(x = g(x)\) by computing \(x_{n+1} = g(x_n)\) with some initial guess \(x_0\). This is called fixed-point iteration.

ex: \(f(x) = x^2 - 3 = 0\)

\[ x = g_1(x) = \frac{3}{x}, \quad x = g_2(x) = x - (x^2 - 3), \quad x = g_3(x) = x - \frac{1}{2}(x^2 - 3) \]

<table>
<thead>
<tr>
<th>(n)</th>
<th>case 1</th>
<th>case 2</th>
<th>case 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(x_n)</td>
<td>(x_n)</td>
<td>(x_n)</td>
</tr>
<tr>
<td>0</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>1</td>
<td>2.00</td>
<td>2.25</td>
<td>1.875</td>
</tr>
<tr>
<td>2</td>
<td>1.50</td>
<td>0.875</td>
<td>1.6172</td>
</tr>
<tr>
<td>3</td>
<td>2.00</td>
<td>3.1523</td>
<td>1.8095</td>
</tr>
<tr>
<td>4</td>
<td>1.50</td>
<td>-3.7849</td>
<td>1.6723</td>
</tr>
</tbody>
</table>
| 5    | 2.00   | -15.1106| 1.7740 | \(\Rightarrow\) Case 1 and case 2 diverge, \(\Rightarrow\) but case 3 converges (recall: \(r = 1.73205\)).
question: what determines whether fixed-point iteration converges or diverges? Consider two examples.

The 1st example diverges and the 2nd example converges.

thm
Assume that $x_0$ is sufficiently close to $r$ and let $k = |g'(r)|$. Then fixed-point iteration converges if and only if $k < 1$.

note: This is consistent with the two examples above.

pf (idea)

| $r - x_{n+1}$ | = $|g(r) - g(x_n)|$ $\sim |g'(r)| \cdot |r - x_n|$ |

Taylor expansion: $g(x_n) = g(r) + g'(r)(x_n - r) + \cdots$

$|r - x_{n+1}| \sim k|r - x_n| \sim k^2|r - x_{n-1}| \sim \cdots \sim k^{n+1}|r - x_0| \quad \text{ok}$

note 1. We showed that $|r - x_{n+1}| \sim k|r - x_n|$. This is called linear convergence and $k$ is called the asymptotic error constant.

recall: $f(x) = x^2 - 3$, $r = \sqrt{3} = 1.73205$

$g_1(x) = \frac{3}{x} \Rightarrow g_1'(x) = -\frac{3}{x^2} \Rightarrow k = |g_1'(r)| = 1 \quad \text{diverges}$

$g_2(x) = x - (x^2 - 3) \Rightarrow g_2'(x) = 1 - 2x \Rightarrow k = |g_2'(r)| = 2.4641 \quad \text{diverges}$

$g_3(x) = x - \frac{1}{2}(x^2 - 3) \Rightarrow g_3'(x) = 1 - x \Rightarrow k = |g_3'(r)| = 0.73205 \quad \text{converges}$

2. The bisection method also converges linearly, with $k = \frac{1}{2}$. 
2.3 Newton’s method

idea: local linear approximation

\[ f(x) \]
\[ \text{tangent line} \]

\[
\text{slope} = f'(x_n) = \frac{0 - f(x_n)}{x_{n+1} - x_n} \Rightarrow x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}
\]

\[
\text{ex}
\]
\[ f(x) = x^2 - 3 \Rightarrow x_{n+1} = x_n - \frac{x_n^2 - 3}{2x_n} \]

| n  | \( x_n \) | \( f(x_n) \) | \( |r - x_n| \) |
|-----|-----------|-------------|-------------|
| 0   | 1.5       | -0.75       | 0.23205081  |
| 1   | 1.75      | 0.0625      | 0.01794919  |
| 2   | 1.73214286| 0.00031888  | 0.00009205  | : rapid convergence |

note

Newton’s method is an example of fixed point iteration, \( x_{n+1} = g(x_n) \), where the iteration function is \( g(x) = x - \frac{f(x)}{f'(x)} \).

\[
g'(x) = 1 - \frac{f'(x)^2 - f(x) \cdot f''(x)}{f'(x)^2} \Rightarrow g'(r) = 1 - \frac{f'(r)^2 - f(r) \cdot f''(r)}{f'(r)^2} = 0
\]

This implies that Newton’s method converges faster than linearly; in fact it can be shown that \( |r - x_{n+1}| \leq C|r - x_n|^2 \), i.e. quadratic convergence.

pf

\[
r - x_{n+1} = g(r) - g(x_n) = g(r) - \left( g(r) + g'(r)(x_n - r) + O((x_n - r)^2) \right) \quad \text{ok}
\]
ex: equation of state of chlorine gas
ideal gas law: \( PV = nRT \), \( P \): pressure, \( V \): volume, \( T \): temperature
\( n \): number of moles present
\( R \): universal gas constant, \( R = 0.08206 \text{ atm} \cdot \text{liter/(mole} \cdot \text{K}) \)

van der Waals equation: \( (P + \frac{n^2a}{V^2})(V - nb) = nRT \)
\( a = 6.29 \text{ atm} \cdot \text{liter}^2/\text{mole}^2 \) (accounts for intermolecular attractive forces)
\( b = 0.0562 \text{ liter/mole} \) (accounts for size of gas molecules)

Take \( n = 1 \) mole, \( P = 2 \) atm, \( T = 313 \) K, and find \( V \) by Newton’s method with starting guess \( V_0 \) given by the ideal gas law.

\[
f(V) = \left( P + \frac{n^2a}{V^2} \right)(V - nb) - nRT, \quad f'(V) = \left( P + \frac{n^2a}{V^2} \right) + \left( \frac{-2n^2a}{V^3} \right)(V - nb)
\]

<table>
<thead>
<tr>
<th>( n )</th>
<th>( V_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>12.84238 99999 99980</td>
</tr>
<tr>
<td>1</td>
<td>12.65115 48134 06302</td>
</tr>
<tr>
<td>2</td>
<td>12.65109 93371 19016 ( \approx 0.2 \text{ atm less than } V_0 ) given by ideal gas law</td>
</tr>
</tbody>
</table>

We infer that \( V_0 \) has 2 correct digits and \( V_1 \) has 5 correct digits. How many correct digits does \( V_2 \) have? (hw)

summary

<table>
<thead>
<tr>
<th>method</th>
<th>rate of convergence</th>
<th>cost per step</th>
</tr>
</thead>
<tbody>
<tr>
<td>bisection</td>
<td>linear, ( k = \frac{1}{2} )</td>
<td>( f(x_n) )</td>
</tr>
<tr>
<td>fixed-point iteration</td>
<td>linear, ( k =</td>
<td>g'(r)</td>
</tr>
<tr>
<td>Newton</td>
<td>quadratic</td>
<td>( f(x_n), f'(x_n) )</td>
</tr>
</tbody>
</table>

note: Bisection is guaranteed to converge if the initial interval contains a root; the other methods are sensitive to the choice of \( x_0 \).

root-finding for nonlinear systems

ex: chemical reactions

\[
\begin{align*}
2A + B & \rightleftharpoons C \\
A + D & \rightleftharpoons C
\end{align*}
\]
reversible reactions for reactants \( A, B, D \) and product \( C \)
\( a_0, b_0, d_0 \): initial concentrations (moles/liter) in chemical reactor (known)
\( c_1, c_2 \): equilibrium concentrations of \( C \) produced by each reaction (unknown)
\( k_1, k_2 \): equilibrium reaction constants (known)
These variables are related by the law of mass action.

\[ \frac{c_1 + c_2}{(a_0 - 2c_1 - c_2)^2(b_0 - c_1)} = k_1 \]

\[ \frac{c_1 + c_2}{(a_0 - 2c_1 - c_2)(d_0 - c_2)} = k_2 \]

Hence to find \( c_1, c_2 \) we need to solve a system of nonlinear equations with 2 equations and 2 unknowns.

**Newton’s method for nonlinear systems**

First note the following alternative derivation of Newton’s method for the case of 1 equation and 1 unknown, \( f(x) = 0 \).

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

Now consider a system of 2 equations and 2 unknowns.

\( f(x, y) = 0, g(x, y) = 0 \)

Given \((x_n, y_n)\), we want to find \((x_{n+1}, y_{n+1})\).

\[ f(x_{n+1}, y_{n+1}) = f(x_n, y_n) + \frac{\partial f}{\partial x}(x_n, y_n)(x_{n+1} - x_n) \]

\[ + \frac{\partial f}{\partial y}(x_n, y_n)(y_{n+1} - y_n) + \cdots \cdots \]

\[ g(x_{n+1}, y_{n+1}) = g(x_n, y_n) + \frac{\partial g}{\partial x}(x_n, y_n)(x_{n+1} - x_n) \]

\[ + \frac{\partial g}{\partial y}(x_n, y_n)(y_{n+1} - y_n) + \cdots \cdots \]

\[ \Rightarrow \begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix}_{(x_n, y_n)} \cdot \begin{pmatrix} x_{n+1} - x_n \\ y_{n+1} - y_n \end{pmatrix} = \begin{pmatrix} -f(x_n, y_n) \\ -g(x_n, y_n) \end{pmatrix} \]

\( \uparrow \)

Jacobian matrix

**note**

1. Given \((x_n, y_n)\), we can solve for \((x_{n+1}, y_{n+1})\). Each step has the form \( Ax = b \), where \( A \) is a given matrix, \( b \) is a given vector, and we must solve for the vector \( x \).

2. hw3 has an application to the chemical reaction system