Seven bar truss structure – Task 1:
Identify Design Variables

Design Parameters
- The cross section area $A_i$
- Material properties of the members

Using Symmetry
Formulate the constraints

- The tensile and compressive stress must not be more than the corresponding allowable strength
- Assume $S_{ty} = S_{cy} = 500\, MPa$
- Compute the axial force generated
  - In AB $\frac{P \csc \theta}{2A_1} \\ _{\leq} S_{yc}$
  - In BC $\frac{P \csc \alpha}{2A_2} \\ _{\leq} S_{yr}$
  - In AC $\frac{P \cos \alpha}{2A_3} \\ _{\leq} S_{yr}$
  - In BD $\frac{P}{2A_4} (\cot \theta + \cot \alpha) \\ _{\leq} S_{yc}$
Stability consideration

- Buckling of the compression members AB, BD, and DE
- Euler buckling conditions

\[
\frac{P}{2 \sin \theta} \leq \frac{\pi E A_l^2}{1.281 l^2}
\]
\[
\frac{P}{2} \left( \cot \theta + \cot \alpha \right) \leq \frac{\pi E A_l^2}{5.76 l^2}
\]

Stiffness Constraint

- Maximum vertical deflection at C \( \delta_{\text{max}} = 2 \text{mm} \)
- Thus

\[
\frac{P l}{E} \left( \frac{0.566}{A_1} + \frac{0.500}{A_2} + \frac{2.236}{A_3} + \frac{2.700}{A_4} \right) \leq \delta_{\text{max}}
\]
Objective Function

- minimizing the weight is equivalent to
- minimization of the total volume of material \( 1.132A_1 + 2A_2 + 1.789A_3 + 1.2A_4 \)

- Hence, Minimize

Bounds of Variables

- Set some lower and upper bounds for the four cross sectional areas

- Say, all four areas lie between 10 and 500 mm\(^2\)

- \(10 \times 10^{-6} \leq A_1, A_2, A_3, A_4, \leq 500 \times 10^{-6}\)
Optimization Problem
Formulation

- Minimize \( \frac{1.132 A_1 l + 2 A_2 l + 1.789 A_3 l + 1.2 A_4 l}{2 A_1 \sin \theta} \geq 0, \)

subject to

\[
\begin{align*}
S_{wc} &= \frac{P}{2 A_1 \sin \theta} \geq 0, \\
S_{s} &= \frac{P}{2 A_2 \cot \theta} \geq 0, \\
S_{s} &= \frac{P}{2 A_3 \sin \alpha} \geq 0, \\
S_{wc} &= \frac{P}{2 A_4} (\cot \theta + \cot \alpha) \geq 0,
\end{align*}
\]

\[
\begin{align*}
\frac{\pi E A_1^2}{1.28 l^2} - \frac{P}{2 \sin \theta} \geq 0, \\
\frac{\pi E A_2^2}{5.76 l^2} - \frac{P}{2} (\cot \theta + \cot \alpha) \geq 0, \\
\delta_{max} &= \frac{P l}{E} \left( \frac{0.566}{A_1} + \frac{0.500}{A_2} + \frac{2.236}{A_3} + \frac{2.700}{A_4} \right) \geq 0,
\end{align*}
\]

\(10 \times 10^{-6} \leq A_1, A_2, A_3, A_4 \leq 500 \times 10^{-6}\)
• Constraints
  - Equality \( \delta(x) = \text{const} \)
  - Inequality \( \sigma(x) \leq \sigma_y \)

• Objective functions
  - Maximize or Minimize
  - The principle of duality can be used to convert maximize problem into minimize problem and vice versa.

Minimize \( f(x) \)
Subject to
\[
\begin{align*}
g_j(x) &\geq 0 & j &= 1, 2, 3 \ldots J \\
h_k(x) &= 0 & k &= 1, 2, 3 \ldots K \\
x^L_i &\leq x_i \leq x^U_i & i &= 1, 2, 3 \ldots N
\end{align*}
\]

Vehicle Suspension design
4 DOF model of car

- Differential equations of the vertical motion of :-
  - The unsprung mass at the front axle \( q_1 \)
    \[
    \ddot{q}_1 = \frac{(F_2 + F_3 - F_1)}{m_{f_1}}
    \]
  - The sprung mass \( q_2 \)
    \[
    \ddot{q}_2 = -\frac{(F_1 + F_3 + F_4 + F_5)}{m_s}
    \]
  - The unsprung mass at the rear axle \( q_3 \)
    \[
    \ddot{q}_3 = \frac{(F_3 + F_5 - F_6)}{m_{r_3}}
    \]
  - The angular motion of the sprung mass \( q_4 \)
    \[
    \ddot{q}_4 = \frac{[(F_4 + F_5)l_2 - (F_2 + F_3)l_1]}{J}
    \]

\( l_1 \) and \( l_2 \) are the horizontal distance of the front and rear axle from the C.G of the sprung mass.
• The forces $F_1$ to $F_6$ are calculated as follows

\[
F_1 = k_f d_1, \quad F_2 = k_f d_2, \quad F_3 = \alpha_f \dot{d}_2, \\
F_4 = k_r d_3, \quad F_5 = \alpha_r \dot{d}_3, \quad F_6 = k_r d_3
\]

• The parameters $d_1, d_2, d_3$ and $d_4$ are the relative deformations in front tyre, the rear spring, the rear tyre and the rear spring respectively. $f_1(t)$ and $f_2(t)$ are road deformations

\[
d_1 = q_1 - f_1(t), \quad d_2 = q_2 + l_1 q_3 - q_1, \\
d_3 = q_4 - f_2(t), \quad d_4 = q_2 - l_2 q_3 - q_4.
\]

• Constraint

say max jerk \( \leq 18 \)

\[
\max \ddot{q}_2(t) \leq 18
\]

• Objective function

Minimize \( \max abs \frac{q_2(t)}{A} \)

Where A = road excitation amplitude

Limits

\[
0 \leq k_f, k_r \leq 2 \text{kg/mm} \\
0 \leq \alpha_f, \alpha_r \leq 300 \text{kg/(m/s)}
\]

Subjected to

\[
18 - \max(\ddot{q}_2(t)) \geq 0
\]
CLASSICAL OPTIMIZATION TECHNIQUES

• SINGLE-VARIABLE OPTIMIZATION

• MULTI-VARIABLE OPTIMIZATION
  - WITH NO CONSTRAINTS
  - WITH EQUALITY CONSTRAINTS
  - WITH INEQUALITY CONSTRAINTS

SINGLE VARIABLE OPTIMIZATION

• FUNCTION HAVING SINGLE VARIABLE

  $$f(x)$$

• FUNCTION FOR THE DIFFERENT VALUES OF (x) CAN HAVE
  - RELATIVE OR LOCAL MINIMUM
  - RELATIVE OR LOCAL MAXIMUM
  - ABSOLUTE OR GLOBAL MINIMUM
  - ABSOLUTE OR GLOBAL MAXIMUM
FUNCTION IS SAID TO HAVE

**- RELATIVE OR LOCAL MINIMUM**
- at \( x = x^* \) if \( f(x^*) \leq f(x^* + h) \)

**- RELATIVE OR LOCAL MAXIMUM**
- at \( x = x^* \) if \( f(x^*) \geq f(x^* + h) \)

For all sufficiently small positive and negative values of \( h \)

---

FUNCTION IS SAID TO HAVE

**- ABSOLUTE OR GLOBAL MINIMUM**
- at \( x = x^* \) if \( f(x^*) \leq f(x) \)

**- ABSOLUTE OR GLOBAL MAXIMUM**
- at \( x = x^* \) if \( f(x^*) \geq f(x) \)

for all \( x \) in the domain defined for \( f(x) \)
Single variable optimization

- Bracketing (exhaustive search)
- Region elimination (internal halving)

**UNIMODAL FUNCTION**

A unimodal function is one that has only one peak (maximum) or valley (minimum) in a given interval.

Mathematically, a function $f(x)$ is unimodal if

(i) $x_1 < x_2 < x^*$ implies that $f(x_2) < f(x_1)$, and

(ii) $x_2 > x_1 > x^*$ implies that $f(x_1) < f(x_2)$, where $x^*$ is the minimum point.
**Exhaustive search algorithm** (given \( f(x) \), \( a \) & \( b \))

**Step 1** set \( x_1 = a \), \( \Delta x = \frac{(b-a)}{n} \) (\( n \) is the number of intermediate points), \( x_2 = x_1 + \Delta x \), and \( x_3 = x_2 + \Delta x \).

**Step 2** If \( f(x_1) \geq f(x_2) \leq f(x_3) \), the minimum point lies in \((x_1, x_3)\), **Terminate**;
Else \( x_i = x_2, x_2 = x_3, x_3 = x_2 + \Delta x \), and go to Step 3.

**Step 3** Is \( x_3 \leq b \) ? If yes, go to Step 2;
Else no minimum exists in \((a,b)\) or a boundary point \((a \text{ or } b)\) is the minimum point.

---

**Interval Halving Method**

**Step 1** Choose a lower bound \( a \) and an upper bound \( b \). Choose also a small number \( \varepsilon \). Let \( x_m = \frac{(a+b)}{2} \), \( L_0 = L = b - a \). Compute \( f(x_m) \).

**Step 2** Set \( x_i = a + L/4 \), \( x_2 = b - L/4 \). Compute \( f(x_1) \) and \( f(x_2) \).

**Step 3** If \( f(x_1) < f(x_m) \) set \( b = x_m, x_m = x_1 \) go to Step 5;
Else go to Step 4.

**Step 4** If \( f(x_2) < f(x_m) \) \{\( a = x_m, x_m = x_2 \); go to step 5\}
Else \{\( a = x_1; b = x_2 \); go to step 5\}.

**Step 5** Calculate \( L = b-a \) If \(|L| < \varepsilon \) **Terminate**;
Else go to Step 2.
EXAMPLE

• Minimize \( f(x) = x^2 + \frac{54}{x} \) in the interval (0,5).

• Step 1: \( \varepsilon = 10^{-3}; \ a=0; \ b=5; \ L_0 = 5; \ x_m = 2.5; \ f(x_m) = 27.85 \).

• Step 2: \( x_1 = 1.25; \ x_2 = 3.75 \)
  \[ f(x_1) = 44.7; \ f(x_2) = 28.4 \]

• Step 3: IS \( f(x_1) < f(x_m) \)? NO.

• Step 4: IS \( f(x_2) < f(x_m) \)? NO.
  hence \([1.25 - 3.75]\) i.e \(a=1.25; \ b=3.75\).

• Step 5: \( L = 2.5; \ a=1.25; b=3.75; \ x_m = 2.5; \)
  \[ x_1 = 1.25 + \frac{2.5}{4} = 1.875; \ x_2 = 3.75 - \frac{2.5}{4} = 3.125; \]
  \[ f(x_1) = 32.3; \ f(x_2) = 27.05 \]

• Step 3 IS \( f(x_1) < f(x_m) \)? NO.

• Step 4 IS \( f(x_2) < f(x_m) \)? YES.
  \(a=2.5; \ b=3.75; \ x_m = 3.125\)

• Step 5 L = 1.25 (3.75 - 2.5)

Iteration continues
Point Estimation Methods

• Instead if using only relative values of two points, use their magnitude also
• Typically fit a curve through a number of known points, and find the minimum

Successive Quadratic Estimation

• A quadratic curve is fitted through three points
• Start with three points $x_1$, $x_2$ and $x_3$
Successive Quadratic Estimation

• A general quadratic function is
  \[ \overline{f}(x) = a_0 + a_1(x - x_1) + a_2(x - x_1)(x - x_2) \]
• If \((f_1, x_1), (f_2, x_2)\) and \((f_3, x_3)\) are known,
  \[ a_0 = f_1 \quad a_1 = (f_1 - f_2)/(x_2 - x_1) \quad a_2 = 1/(x_2 - x_1)(x_3 - x_1) \]
• Minima lies at
  \[ x = \frac{x_1 + x_3}{2} - \frac{a_1}{2a_2} \]

Algorithm

• S1: Let \(x_1\) be the initial point, \(\Delta\) be the step size, \(x_2 = x_1 + \Delta\)
• S2: Compute \(f(x_1)\) and \(f(x_2)\)
• S3: If \(f(x_1) > f(x_2)\) Let \(x_3 = x_1 + 2\Delta\) else \(x_3 = x_1 - \Delta\), Compute \(f(x_3)\)
• S4: Determine \(f_{\text{min}} = \min(f_1, f_2, f_3)\) and \(X_{\text{min}}\)
• S5: Compute \(\bar{x}\)
• S 6: If \(\left| f_{\text{min}} - f(\bar{x}) \right|\) and \(\left| x_{\text{min}} - \bar{x} \right|\) are small, optima is best of all known points
• S7: Save the best point and neighbouring points. Goto 4.
Bracketing Method based on unimodal property of objective function

1) Assume an interval \([a,b]\) with a minima in the range
2) Consider \(x_1\) and \(x_2\) within the interval.
3) Find out values of \(f\) at \(x=x_1, x_2\) and Compare \(f(x_1)\) and \(f(x_2)\)
   a) If \(f(x_1) < f(x_2)\) then eliminate \([x_2,b]\) and set new interval \([a,x_2]\)
   b) If \(f(x_1) > f(x_2)\) then eliminate \([a,x_1]\) and set new interval \([x_1,b]\)
   c) If \(f(x_1) = f(x_2)\) then eliminate \([a,x_1]\) & \([x_2,b]\) and set new interval \([x_1,x_2]\)

**FIBONACCI METHOD**

**APPLICATION**

- To find minimum of a function of one variable even if function is not continuous

**LIMITATIONS:**

- The initial interval of uncertainty i.e. range in which the optimum lies, has to be known.
- The function being optimized has to be unimodal in the initial interval of uncertainty.
- The exact optimum cannot be located in this method. Only an interval known as the
  Final Interval of uncertainty will be known. The final interval of uncertainty can be made
Advantage of Fibonacci Series

• Fibonacci Series: \( F_0 = F_1 = 1; \) \( F_n = F_{n-1} + F_{n-2} \)
• Hence: 1, 1, 2, 3, 5, 8, 13, 21, 34, ...

• \( L_k - L_k^* = L_{k+1}^* \)
• Hence one point is always pre-calculated

Fibonacci Search Algorithm

• Step 1: \( L = b - a; \) \( k = 2; \) Decide \( n; \)
• Step 2: 
  \[ L' = \left( \frac{F_{n+k}}{F_{n+k}} \right) L \]  
  \( x_1 = a + L' \)  
  \( x_2 = b - L' \)
• Step 3:
  – Compute either \( f(x_1) \) or \( f(x_2) \) (whichever is not computed earlier)
  – Use region elimination rule
  – Set new \( a \) and \( b \)
• If \( k = n \), TERMINATE else \( k = k + 1 \) and GOTO Step 2
Sequence of Fibonacci numbers is defined as:
\[ F_0 = F_1 = 1; \quad F_n = F_{n-1} + F_{n-2} \quad \text{i.e. 1,1,2,3,5,8,13,21,34,} \ldots \]

- Let initial interval \( I_0 \) defined by \([a, b]\) be known.
- Each function evaluation is termed as an experiment
\[ L_0 = b - a \quad L_2 = \frac{F_{n-2}}{F_n} L_0 \]
- Define
  \[ x_1 = a + L_2 = a + \frac{F_{n-2}}{F_n} L_0 \]
  \[ x_2 = b - L_2 = b - \frac{F_{n-2}}{F_n} L_0 = a + \frac{F_{n-1}}{F_n} L_0 \]
- Discard part of interval using bracketing method for unimodal functions
- Remaining interval to be used in next iteration is
\[ L_2 = L_0 - L_2 = L_0 \left( 1 - \frac{F_{n-2}}{F_n} \right) = \frac{F_{n-1}}{F_n} L_0 \]
- New search interval obtained contains one previous exp. point at a distance \( L_2^* \) from one end (old end) \( L_2^* = \frac{F_{n-2}}{F_n} L_0 \) or \( L_2^* = \frac{F_{n-3}}{F_n} L_0 = \frac{F_{n-3}}{F_{n-1}} L_2 \) from new end

- Place third experiment point in the interval of \( L_2 \)
- Current two experiment points are located at a distance of \( L_3^* \)
\[ L_3^* = \frac{F_{n-3}}{F_n} L_0 = \frac{F_{n-3}}{F_{n-1}} L_2 \]
- Again discard interval based on unimodal property
- New interval of search obtained in \( L_3 \) where
\[ L_3 = L_2^* - L_2 = \frac{F_{n-3}}{F_{n-1}} L_2 = \frac{F_{n-2}}{F_n} L_0 \]
- For any \( j \)th experiment out of \( n \) experiment
\[ L_j = \frac{F_n - (j-2)}{F_n - (j-1)} L_{j-1} \]
- The ratio \( \frac{L_n}{L_0} = \frac{F_1}{F_n} = \frac{1}{\varphi} \) will permit us to determine ’\( n \)’ the required number of experiments to achieve the desired accuracy in locating the optimum point.

- After conducting \( n-1 \) experiments the remaining interval will contain one experiment point precisely at its middle point. The \( n^{th} \) experiment point has also to be placed there. Since no new information can be obtained by placing point there, the nth experiment point is placed very close to the remaining valid experiment point. This enables to obtain final interval of uncertainty to within \( 0.5\% \).
**Golden Section Method**

**APPLICATION**
- To find minimum of a function of one variable even if function is not continuous

The golden section method is same s the Fibonacci method except that in Fibonacci method
1. the ratio is not constant in every iteration
2. total number of experiments to be conducted are specified in the beginning whereas in the golden section method it is not required

The procedure is same as the Fibonacci method except that the location of the first two experiments is given by \( L_2^* = L_0/(\phi)^2 = 0.382* L_0 \)

The desired accuracy can be specified to stop the procedure.
Historical background of Golden Section

The value phi has historical background. It was believed by ancient Greeks that a building having sides d, b such that (d+b)/d = d/b = phi will be having most pleasing properties. Or 1+ (1/phi) = phi

This gives phi = (1+ 50.5)/2 = 1.618

It is also found in Euclid geometry that the division of line into two unequal parts so that the ratio of whole to the larger part equals the ratio of larger to smaller part, being known as the golden section or golden mean.

• In every iteration if you reduce range by a factor r
• Original range – AD; reduced to AC in 1st iteration (points B and C are already calculated)
• If the next range in iteration 2 has to be AB, AB = r*AC;
• HENCE 1−r = r2

Golden Section Search Algorithm

• Step 1: L=b-a; k = 1; Decide ε; map a, b to a’=0; b’=1
• Step 2: L_w=b’-a’
  – w_1=a’+0.618L_w
  – w_2=b’-0.618L_w
• Step 3:
  – Compute either f(w_1) or f(w_2) (whichever is not computed earlier)
  – Use region elimination rule
  – Set new a’ and b’
• If |L_w|≤ε TERMINATE else k=k+1 and GOTO Step 2
### An Example - Golden Section

Function $f(x)$ is

$$0.65 - \frac{0.75}{(1+x^2)} - 0.65 \times \tan^{-1} \frac{1}{x}$$

is minimized using golden section method with $n=6$

$A=0 \quad B=3$

- The location of first two exp. Points are defined by $L_2^* = 0.382L_0 = (0.382)(3.0) = 1.146$

$X_1 = 1.1460 \quad X_2 = 3.0 - 1.1460 = 1.854$

with $f_1 = -0.208 \quad f_2 = -0.115$

Since $f_1 < f_2$ delete $[x_2, 3.0]$ based on assumption of unimodality and new interval of uncertainty obtained is $[0, x_2] = [0, 1.854]$

- The third experiment point is placed at $x_3 = 0 + (x_2 - x_1) = 1.854 - 1.146 = 0.708$.

$f_3 = -0.288943 \quad f_1 = -0.208$

Since $f_3 < f_1$ delete interval $[x_1, x_2]$.

The new interval of uncertainty is $[0, x_1] = [0, 1.146]$

- The fourth experiment point is placed at $x_4 = 0 + (x_1 - x_3) = 0.438$.

$f_4 = -0.308951 \quad f_3 = -0.288943$

Since $f_4 < f_3$ delete interval $[x_3, x_1]$.

The new interval of uncertainty is $[0, x_3] = [0, 0.7080]$

- The fifth experiment point is placed at $x_5 = 0 + (x_3 - x_4) = 0.27$.

$f_4 = -0.308951 \quad f_5 = -0.278$

Since $f_4 < f_5$ delete interval $[0, x_5]$.

The new interval of uncertainty is $[x_5, x_3] = [0.27, 0.7080]$

- The last experiment point is placed at $x_6 = x_5 + (x_3 - x_4) = 0.54$.

$f_4 = -0.308951 \quad f_6 = -0.308234$

Since $f_4 < f_6$ delete interval $[x_6, x_3]$.

The new interval of uncertainty is $[x_5, x_6] = [0.27, 0.54]$

Ratio of final interval to initial interval is $0.09$. 

---

### An Example - Golden Section

- The fourth experiment point is placed at $x_4 = 0 + (x_1 - x_3) = 0.438$.

$f_4 = -0.308951 \quad f_3 = -0.288943$

Since $f_4 < f_3$ delete interval $[x_3, x_1]$.

The new interval of uncertainty is $[0, x_3] = [0, 0.7080]$

- The fifth experiment point is placed at $x_5 = 0 + (x_3 - x_4) = 0.27$.

$f_4 = -0.308951 \quad f_5 = -0.278$

Since $f_4 < f_5$ delete interval $[0, x_5]$.

The new interval of uncertainty is $[x_5, x_3] = [0.27, 0.7080]$

- The last experiment point is placed at $x_6 = x_5 + (x_3 - x_4) = 0.54$.

$f_4 = -0.308951 \quad f_6 = -0.308234$

Since $f_4 < f_6$ delete interval $[x_6, x_3]$.

The new interval of uncertainty is $[x_5, x_6] = [0.27, 0.54]$

Ratio of final interval to initial interval is $0.09$. 

---
### Bounding Phase Method (for bracketing the optima)

<table>
<thead>
<tr>
<th>Step 1</th>
<th>Choose an initial guess $X_0$ and an increment $D$. Set $K = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 2</td>
<td>If $f(X_0 -</td>
</tr>
<tr>
<td></td>
<td>$f(X_0 -</td>
</tr>
<tr>
<td></td>
<td>else goto step 1</td>
</tr>
<tr>
<td>Step 3</td>
<td>Set $X_{K+1} = X_K + 2^K*D$</td>
</tr>
<tr>
<td>Step 4</td>
<td>If $f(X_{K+1}) &lt; f(X_K)$ Set $K = K+1$ and goto step 3</td>
</tr>
<tr>
<td></td>
<td>else, the minimum lies in the interval $(X_{K-1}, X_{K+1})$ and terminate</td>
</tr>
</tbody>
</table>

If $D$ is large, accuracy is poor.

### Bounding Phase Method – An example

Minimize $f(x) = x^2 + 54/x$ ;

<table>
<thead>
<tr>
<th>Step 1</th>
<th>Choose an initial guess $X_0 = 0.6$ and an increment $D = 0.5$. Set $K = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 2</td>
<td>Calculate $f(X_0 -</td>
</tr>
<tr>
<td></td>
<td>$f(X_0 +</td>
</tr>
<tr>
<td>Step 3</td>
<td>Set $X_1 = X_0 + 2^0*D \Rightarrow X_1 = 1.1$</td>
</tr>
<tr>
<td>Step 4</td>
<td>If $f(X_1) = 50.301 &lt; f(X_0)$. Set $K = 1$ and goto step 3</td>
</tr>
<tr>
<td></td>
<td>Next guess $X_2 = X_1 + 2^1*D = 2.1$</td>
</tr>
<tr>
<td></td>
<td>$*f(X_2) = 30.124 &lt; f(X_1)$ therefore set $K=2$ and goto step 3</td>
</tr>
<tr>
<td></td>
<td>$<em>\text{Next guess } X_3 = X_2 + 2^2</em>D = 4.1$</td>
</tr>
<tr>
<td></td>
<td>$*f(X_3) = 29.981 &lt; f(X_2)$ therefore set $K=3$ and goto step 3</td>
</tr>
<tr>
<td></td>
<td>$<em>\text{Next guess } X_4 = X_3 + 2^3</em>D \quad f(X_4) = 72.277 &gt; f(X_3)$.</td>
</tr>
<tr>
<td></td>
<td>$*\text{Thus terminate with interval (2.1,8.1)}$</td>
</tr>
</tbody>
</table>

With $D = 0.5$ the bracketing is poor.

If $D = 0.001$, the obtained interval is $(1.623, 4.695)$. 
Mathematical techniques and gradient based approaches

Theorem 2.1: **Necessary condition**

- If a function $f(x)$ is defined in the interval $a \leq x \leq b$ and has a relative minimum at $x = x^*$, where $a < x^* < b$, and if the derivative $df(x)/dx = f'(x)$ exists as a finite number at $x = x^*$, then $f'(x^*) = 0$
Proof:

- It is given that

\[ f'(x) = \lim_{h \to 0} \frac{f(x^* + h) - f(x^*)}{h} \] ..................1

Exists as a definite number, which we want to prove to be zero. Since \( x^* \) is a relative minimum, we have

\[ f \leq f(x^* + h) \]

for all values of \( h \) sufficiently close to zero. Hence

\[ \frac{f(x^* + h) - f(x^*)}{h} \geq 0 \quad \text{if} \quad h > 0 \]

\[ \frac{f(x^* + h) - f(x^*)}{h} \leq 0 \quad \text{if} \quad h < 0 \]

Thus equation 1 gives the limit as \( h \) tends to zero through positive values as

\[ f''(x^*) \geq 0 \]

while 1 gives the limit as \( h \) tends to zero through negative values as

\[ f''(x^*) \leq 0 \]

The only way to satisfy both the equations above is to have

\[ f''(x^*) = 0 \]

This proves the theorem
Theorem 2.2: Sufficient Condition

- Let $f''(x^*) = f'''(x^*) = \ldots = f^{(n+1)}(x^*)$, $(x^*)=0$, but $f^{(n)}(x^*) \neq 0$
- Then $f(x)$ is
  
  (i) a minimum value of $f(x)$ if $f^{(n)}(x^*)>0$ and $n$ is even;
  
  (ii) a maximum value of $f(x)$ if $f^{(n)}(x^*)<0$ and $n$ is even;
  
  (iii) neither a maximum nor a minimum if $n$ is odd.

Proof: Applying Taylor’s theorem with reminder after $n$ terms, we have

$$f(x^*+h) = (f(x^*) + hf'(x^*) + \frac{h^2}{2!}f''(x^*) + \ldots + \frac{h^{n-1}}{(n-1)!}f^{(n-1)}(x^*) + \frac{h^n}{n!}f^{(n)}(x^*+\theta h)) \quad for \quad 0 < \theta < 1$$

Since $f'(x^*) = f''(x^*) = \ldots = f^{(n-1)}(x^*) = 0$

Hence the above equation becomes

$$f(x^*+h) - f(x^*) = \frac{h^n}{(n)!}f^{(n)}(x^*+\theta h)$$
• As $f^{(n)}(x^*) \neq 0$, there exist an interval around $x^*$ for every point of $x$ of which the $n^{th}$ derivative $f^{(n)}(x^*)$ has the same sign, namely, that of $f^{(n)}(x^*)$. Thus for every point $x^*+h$ of this interval $f^{(n)}(x^*+\theta h)$ has the sign of $f^{(n)}(x^*)$

When $n$ is even, $h^n/n!$ is positive irrespective of whether $h$ is positive or negative, and hence $f^{(n)}(x^*+h)-f(x^*)$ will have the same sign as that of $f^{(n)}(x^*)$

Thus $x^*$ will be
- Relative minimum if $f^{(n)}(x^*)$ is positive
- Relative maximum if $f^{(n)}(x^*)$ is negative

When $n$ is odd $h^n/n!$ changes sign with the change in the sign of $h$ and hence the point $x^*$ is neither maximum nor a minimum.
In this case point $x^*$ is called a point of inflection
Determine the maximum and minimum values of the function \( f(x) = 12x^5 - 45x^4 + 40x^3 + 5 \)

Taking the first derivative of the function yields to ...

\[ f'(x) = 60x^4 - 180x^3 + 120x^2 \]

The value of the function at \( x = 0 \), \( x = 1 \), \( x = 2 \) is zero

In the next step take the second derivative of the function
Second derivative of the function yields to ...

\[ f''(x) = 240x^3 - 630x^2 + 240x \]

The value of the function at
\[ x = 1, f''(x) = -60 \] hence \( x = 1 \) is a relative maximum
\[ f_{\text{max}} \mid_{x=1} = 12 \]
\[ x = 2, f''(x) = 240 \] hence \( x = 2 \) is a relative minimum
\[ f_{\text{min}} \mid_{x=2} = -11 \]
\[ x = 0, f''(x) = 0 \] hence evaluate the next derivative

Evaluate the third derivative

Third derivative of the function yields to ...

\[ f'''(x) = 60(12x^2 - 18x + 4) = 240 \] at \( x = 0 \)

Since \( f'''(x) \neq 0 \) at \( x=0, x=0 \) is neither a maximum nor a minimum, and it is an inflection point
In a two stage compressor, the working gas leaving the first stage of compression is cooled (by passing it through a heat exchanger) before it enters the second stage of compression to increase the efficiency. The total work input to a compressor \( W \) for isentropic compression, is given by

\[
W = c_p T_1 \left[ \left( \frac{p_2}{p_1} \right)^{(k-1)/k} + \left( \frac{p_3}{p_2} \right)^{(k-1)/k} - 2 \right]
\]

Where \( c_p \) is the specific heat of the gas at constant pressure, \( k \) is the ratio of specific heat at constant pressure to that at constant volume of the gas, and \( T_1 \) is the temperature at which the gas enters the compressor. Find the pressure, \( p_2 \), at which inter-cooling should be done to minimize the work input to the compressor. Also determine the minimum work done on the compressor.

The necessary condition for minimizing the work done on the compressor is,

\[
\frac{dW}{dp_2} = c_p T_1 k \left[ k-1 \left( \frac{1}{p_1} \right)^{(k-1)/k} \frac{k-1}{k} (p_2)^{-1/k} + p_3 \frac{(k-1)/k}{k} - k+1 \right] = 0
\]

Which yields

\[
p_2 = \left( p_1 p_2 \right)^{1/2}
\]

The second derivative yields

\[
\frac{d^2W}{dp_2^2} = c_p T_1 k \left[ - \left( \frac{1}{p_1} \right)^{(k-1)/k} \frac{1}{k} (p_2)^{-1/k} - p_3 \frac{(k-1)/k}{k} - 2k+1 \right]
\]
\[
\left( \frac{d^2W}{dp_2^2} \right)_{p_2=(p_1p_3)^{1/2}} = \frac{2c_p T_1 \frac{k-1}{k}}{P_1^{(3k-1)/2k} P_3^{(k+1)/2k}}
\]

Since the ratio of specific heats \( k \) is greater than 1, we get

\[
\frac{d^2W}{dp_2^2} > 0 \quad \text{at} \quad p_2 = (p_1p_3)^{1/2}
\]

Hence the solution corresponds to relative minimum.

The minimum work done is given by

\[
W_{\text{min}} = c_p T_1 \frac{k}{k-1} \left[ \left( \frac{P_3}{P_1} \right)^{(k-1)/2k} - 1 \right]
\]
Gradient based Methods

- Algorithms require derivative information
- Many real world problems, difficult to obtain information about derivatives
  - Computations involved
  - Nature of problem
- Still gradient methods are effective and popular
- Recommended to use in problems where derivative information available
- **Global optimum** occurs where gradient is zero
- the search process terminates where gradient is zero

Overview of methods

- Methods
  - Newton-Raphson method
  - Bisection method
  - Secant method
  - Cubic search method
Newton-Raphson Method

- Goal of unconstrained optimization – reach as small a derivative as possible
- A linear approximation of first derivative of function at a point is expressed using Taylor’s series expansion.
- Equate to zero to find next guess
- If the current point at iteration $t$ is $x^{(t)}$, the next iteration is governed by following expression

$$x^{(t+1)} = x^{(t)} - \frac{f'(x^{(t)})}{f''(x^{(t)})}$$

Newton-Raphson Method

Algorithm

**Step 1**: Choose initial guess $x^{(1)}$ and a small number $\varepsilon$. Set $k=1$. Compute $f'(x^{(1)})$

**Step 2**: Compute $f''(x^{(k)})$

**Step 3**: Calculate

$$x^{(k+1)} = x^{(k)} - \frac{f'(x^{(k)})}{f''(x^{(k)})} \quad (1)$$

**Step 4**: If $|f'(x^{(k+1)})| < \varepsilon$, Terminate;
Else set $k = k+1$ and go to **Step 2**
Newton Raphson Method

- Convergence depends on the initial point and the nature of objective function
- In Practice, gradients have to be computed numerically
- At a point $x^{(t)}$, using central difference method:

\[
\begin{align*}
    f'(x^{(t)}) &= \frac{f(x^{(t)} + \Delta x^{(t)}) - f(x^{(t)} - \Delta x^{(t)})}{2\Delta x^{(t)}} \quad (2) \\
    f''(x^{(t)}) &= \frac{f(x^{(t)} + \Delta x^{(t)}) - 2f(x^{(t)}) + f(x^{(t)} - \Delta x^{(t)})}{(\Delta x^{(t)})^2} \quad (3)
\end{align*}
\]

Newton Raphson Method

The parameter $\Delta x^{(t)}$ is usually taken to be small

\[
\Delta x^{(t)} = \begin{cases} 
0.01 & \text{if } |x^{(t)}| > 0.01 \\
0.01 & \text{otherwise}
\end{cases} \quad (4)
\]

The first derivative requires two function evaluations and the second derivative requires three function evaluations
Newton-Raphson Method

Consider minimization problem

\[ f(x) = x^2 + 54/x \]

**Step 1:** We choose an initial guess \( x^{(1)} = 1 \), a termination factor of \( \varepsilon = 10^{-3} \), and an iteration counter \( k = 1 \). Computing derivative using (2). The small increment using (4) is 0.01. The computed derivative is -52.005, whereas the exact derivative at \( x^{(1)} \) is found to be -52. It is accepted and proceeding to **Step 2**

Newton-Raphson Method

**Step 2:** The exact second derivative of function \( x(1) = 1 \) is found to be 110. The second derivative computed using (3) is 110.011, which is close to exact value.

**Step 3:** We compute the next guess,

\[
x^{(2)} = x^{(1)} - \frac{f'(x^{(1)})}{f''(x^{(1)})}
\]

\[
= 1 - \left( -52.005 \right) / (110.011 )
\]

\[
= 1.473
\]
Newton-Raphson Method

The derivative computed using (2) at this point is found to be \( f'(x^{(2)}) = -21.944 \)

**Step 4**: Since \( |f'(x^{(2)})| \) not less than \( \varepsilon \), we increment \( k \) to 2 and go to **Step 2**. This completes one iteration of the Newton-Raphson method.

**Step 2**: (Second Iteration) \( f''(x^{(2)}) = 35.796 \)

**Step 3**: (Second Iteration) The next guess \( x(3) = 2.086 \) and \( f'(x^{(3)}) = -8.239 \)

**Step 4**: (Second Iteration) Since \( |f'(x^{(3)})| \) not less than \( \varepsilon \), we increment \( k \) to 3 and go to **Step 2**.

---

Newton-Raphson Method

**Step 2**: (Third Iteration) The second derivative at the point \( f''(x^{(3)}) = 13.899 \)

**Step 3**: (Third Iteration) The new point is calculated as \( f'(x^{(4)}) = -2.167 \). (Nine function evaluations)

**Step 4**: (Third Iteration) Since the absolute value of this derivative not smaller than \( \varepsilon \), the search proceeds to **Step 2**.

After three more iterations \( x^{(7)} = 3.0001 \) and \( f'(x^{(7)}) = -4(10)^{-8} \), small enough to terminate algorithm

Since at every iteration first and second order derivatives are evaluated, a total of three function values are evaluated for every iteration
Bisection Method

• Computation of second derivative is avoided; only first derivative is used.
• Both function and the value and the sign of the first derivative at two points is used to eliminate a certain portion of the search space.
• Method similar to the region-elimination methods discussed.
• Algorithm assumes unimodality of the function.

Bisection Method

• Using derivative information, the minimum, is said to be bracketed in the interval (a,b) if two conditions – \( f'(a) < 0 \) and \( f'(b) > 0 \) are satisfied.
• Requires two initial boundary points bracketing the minimum.
• Derivatives at two boundary points and at the middle point are calculated and compared.
• Of three points, two consecutive points with derivatives having opposite signs are chosen for next iteration.
Bisection Method

Algorithm:

**Step 1**: Choose two points \(a\) and \(b\) such that \(f'(a)<0\) and \(f'(b)>0\). Also choose a small number \(\varepsilon\). Set \(x_1 = a\) and \(x_2 = b\).

**Step 2**: Calculate \(z=(x_2+x_1)/2\) and evaluate \(f'(z)\)

**Step 3**: If \(|f'(z)| \leq \varepsilon\), TERMINATE;
Else if \(f'(z) < 0\) set \(x_1 = z\) and go to **Step 2**
Else if \(f'(z) > 0\) set \(x_2 = z\) and go to **Step 2**

Bisection Method

- The sign of first-derivative at the mid-point of the current search region is used to eliminate half of the search region.
  - If derivative –ve, minimum cannot lie in left half of search region
  - If derivative +ve, the minimum cannot lie on the right half of the search space.
Bisection Method

Consider again the function:

\[ f(x) = x^2 + 54/x \]

**Step 1:** Choose two points \( a = 2 \) and \( b = 5 \) such that \( f'(a) = -9.501 \) and \( f'(b) = 7.841 \) are of opposite sign. \( \varepsilon = 10^{-3} \)

**Step 2:** Calculate a quantity \( z = (x_1 + x_2)/2 = 3.5 \) and compute \( f'(z) = 2.591 \).

**Step 3:** Since \( f'(z) > 0 \), the right half of the search space is eliminated. \( x_1 \) is set as 2 and \( x_2 = x = 3.5 \) thus one iteration completed.

At each iteration, one half of the search region is eliminated but here the decision about which half to eliminate depends on the derivatives at the mid point of the interval.

Bisection Method

**Step 2:** (Second Iteration) \( z = (2 + 3.5)/2 = 2.750 \) and \( f'(z) = -1.641 \).

**Step 3:** (Second Iteration) since \( f'(z) < 0 \), \( x_1 = 2.750 \) and \( x_2 = 3.50 \).

**Step 2:** (Third Iteration) \( z = 3.125 \) and \( f'(z) = 0.720 \)

**Step 3:** (Third Iteration) Since \( |f'(z)| \) not less than \( \varepsilon \), iteration continued.

At the end of 10 function evaluations intervals \((2.750, 3.125)\), bracketing of minimum point \( x^* = 3.0 \). The guess of minimum point is obtained at interval of \( x = 2.938 \). This process continues until we find a vanishing derivative.
Secant Method

- Both magnitude and sign of derivatives to create a new point
- Derivative of function assumed to vary linearly between two chosen boundary points.
- If two points \( x_1 \) and \( x_2 \), the quantity \( f'(x_1) f'(x_2) \leq 0 \), the linear approximation of the derivative \( x_1 \) and \( x_2 \) will have a zero derivative at a point \( z \) given by

\[
z = x_2 - \frac{f''(x_2)}{(f''(x_2) - f''(x_1))/(x_2 - x_1)} \tag{5}
\]

Secant Method

- In this method, in one iteration, more than half the search space may be eliminated depending on the gradient values at the chosen points.

Algorithm:

**Step 1**: Algorithm same as bisection method except that **Step 2**

**Step 2**: Calculate new point \( z \) using (5) and evaluate \( f'(z) \).

This algorithm also requires only one gradient evaluation at every iteration. Thus, only two function values are required per iteration.
Secant Method

Considering the problem
\[ f'(x) = x^2 + 54/x \]

**Step 1:** Initial points \( a=2, \ b=5 \) having derivatives
\[ f'(a) = -9.501 \] and \[ f'(b) = 7.841 \] with opposite signs. Choosing \( \epsilon = 10^{-3} \) and set \( x_1 = 2 \) and \( x_2 = 5 \).

**Step 2:** Calculating new point using equation (5):
\[
z = 5 - \frac{f'(5)}{(f'(5) - f'(2))/(5 - 2)} = 3.644
\]

Secant Method

**Step 3:** Since \( f'(z) > 0 \), we eliminate the right part (the region \((z,b))\) of the original search region. The amount of eliminated search space is \( (b-z) = 1.356 \), which is less than half the search space \( (b-a)/2 = 2.5 \). Here, \( x_1 = 2 \) and \( x_2 = 3.644 \) and proceed with next iteration.

**Step 2:** (Second Iteration) \( z = 3.228 \) and \( f'(z) = 1.127 \)

**Step 3:** (Second Iteration) Right part of search space eliminated since \( f'(z) > 0 \). \( x_1 = 2 \) and \( x_2 = 3.228 \) for next iteration
Secant Method

Step 2: (Third Iteration) \( z = 3.101 \) and \( f'(z) = 0.586 \)

Step 3: (Third Iteration) since \( |f'(z)| \) not less than \( \varepsilon \), next iteration proceeded.

At end of 10 function evaluations, true minimum is computed \( x=3.037 \). It is closer to bisection method.

Cubic Search Method

- Similar to successive quadratic point-estimation method except that the derivatives are used to reduce the number of required initial points. For ex.

  \[
  \tilde{f}(x) = a_0 + a_1(x-x_1) + a_2(x-x_1)(x-x_2) + a_3(x-x_1)^2(x-x_2)
  \]

- \( a_0, a_1, a_2, a_3 \) be unknowns therefore requires at least four points to determine the function.

- Function can also be determined by specifying the function value as well as derivative at two points
Cubic Search Method

• \(((x_1, f_1, f'_1), (x_2, f_2, f'_2))\) and by setting equation to zero.

\[
\bar{x} = \begin{cases} 
  x_2, & \text{if } \mu = 0 \\
  x_2 - \mu(x_2 - x_1), & \text{if } 0 \leq \mu \leq 1 \\
  x_1, & \text{if } \mu > 1
\end{cases}
\]

(6)

• Where

\[
z = \frac{3(f_1 - f_2)}{x_2 - x_1} + f'_1 + f'_2
\]

\[
w = \frac{x_2 - x_1}{|x_2 - x_1|}\sqrt{(z^2 - f'_1 f'_2)}
\]

\[
\mu = \frac{f'_2 + w - z}{f'_2 - f'_1 + 2w}
\]

Cubic Search Method

• Similar to Powell’s successive quadratic estimation method, estimation of \(f(x)\) can be used to estimate true minimum of objective function.

• Estimate with earlier two points are used to find the next estimate of the true minimum point.

• The product of their derivative of two points is negative.

• This procedure is continued till desired accuracy
Cubic Search Method

Algorithm:

**Step 1:**
- Choose initial point \( x^{(0)} \), a step size \( \Delta \) and two termination parameters \( \varepsilon_1 \) and \( \varepsilon_2 \).
- Compute \( f'(x^{(0)}) \).
- If \( f'(x^{(0)}) > 0 \), set \( \Delta = -\Delta \). Set \( k = 0 \).

**Step 2:** Compute \( x^{(k+1)} = x^{(k)} + 2^k \Delta \)

**Step 3:** Evaluate \( f'(x^{(k+1)}) \)
- If \( f'(x^{(k+1)}) f'(x^{(k)}) \leq 0 \), set \( x_1 = x^{(k)} \), \( x_2 = x^{(k+1)} \), and go to **Step 4**
- Else set \( k = k+1 \) and go to **Step 2**.

**Step 4:** Calculate the point using (6)

**Step 5:** If \( f(x) \leq f(x_{1}) \), go to **Step 6**
Else set \( \bar{x} = \bar{x} - ((x - x_{1})/2) \) until \( f(\bar{x}) \leq f(x_{1}) \) is achieved

**Step 6:** Compute \( f'(\bar{x}) \). If \( |f'(\bar{x})| \leq \varepsilon_1 \) and \( |(\bar{x} - x_{1})/\bar{x}| \leq \varepsilon_2 \), Terminate
Else if \( f'(\bar{x})f'(x_{1}) < 0 \), set \( x_{2} = \bar{x} ; \)
Else set \( x_{1} = \bar{x} \)
Go to **Step 4**
Cubic Search Method

- Method is most effective if the exact derivative is available
- Bracketing of minimum point achieved in first three steps
- Bracketing algorithm similar to bounding phase method
- Except first iteration, function value as well as first derivative are calculated only at one new point ➔ Only two new function evaluations required.
- First iteration requires repetitive execution of steps 2 and 3 to obtain bracketing points.

Cubic Search Method

- If new point $\dot{x}$ is better than $x_1$, one of the two points is eliminated depending on which brackets the true minimum with $\ddot{x}$.
- If new point is worse than $x_1$, the best two among the points and $\ddot{x}$ will bracket.
- Excessive derivative computations are avoided by simply modifying the point $\ddot{x}$. 
Cubic Search Method

Considering the function
\[ f'(x) = x^2 + 54 / x \]

**Step 1:**
- Choosing an initial point \( x^{(0)} = 1 \), a step size \( \Delta = 0.5 \) and termination parameters \( \varepsilon_1 = \varepsilon_2 = 10^{-3} \).
- The derivative of the function at \( x^{(0)}, f'(x^{(0)}) = -52.005 \)
- \( f'(x^{(0)}) < 0 \), hence \( \Delta = 0.5, k=0 \)

**Step 2:**
- \( x^{(1)} = x^{(0)} + 2(0) \Delta = 1 + 1(0.5) = 1.5 \)
- \( f'(x^{(1)}) = -21.002 \)
- Hence, \( f'(x^{(0)})f'(x^{(1)}) \) not less than or equal to 0

**Step 3:**
- Derivative \( f'(x^{(2)}) = -3.641 \) does not make product negative.
- \( x^{(3)} = 1.5 + 2^2(0.5) = 4.5 \) \( \Rightarrow f'(x^{(3)}) = 6.333 \), which makes the product \( f'(x^{(2)})f'(x^{(3)}) \) negative
- \( x_1 = x^{(2)} = 2.5 \) and \( x_2 = x^{(3)} = 4.5 \)

Cubic Search Method

- \( x^{(1)} = x^{(0)} + 2^{(0)} \Delta = 1 + 1(0.5) = 1.5 \)
- \( f'(x^{(1)}) = -21.002 \)
- Hence, \( f'(x^{(0)})f'(x^{(1)}) \) not less than or equal to 0

**Step 2:** Hence, \( k = 1 \) and
- \( x^{(2)} = 1.5 + 2^1(0.5) = 2.5 \)

**Step 3:**
- Derivative \( f'(x^{(2)}) = -3.641 \) does not make product negative.
- \( x^{(3)} = 1.5 + 2^2(0.5) = 4.5 \) \( \Rightarrow f'(x^{(3)}) = 6.333 \), which makes the product \( f'(x^{(2)})f'(x^{(3)}) \) negative
- \( x_1 = x^{(2)} = 2.5 \) and \( x_2 = x^{(3)} = 4.5 \)
Cubic Search Method

**Step 4:** Estimated point is calculated
- \( z = -3.907 \), \( w = 6.190 \) and \( \mu = 0.735 \).
- \( x = 4.5 - 0.735(4.5 - 2.5) = 3.030 \)

**Step 5:**
- \( f(x) = 27.003 \)
- Since \( f(x) < f(x_1) \), go to **Step 6**.

**Step 6:**
- \( f'(x) = 0.178 \) ➔ Termination criteria not satisfied
- \( f'(x)f'(x_1) < 0 \), hence \( x_2 = x = 3.030 \)
- End of iteration one. and go to **Step 4**

---

Cubic Search Method

**Step 4:** (Second Iteration)
- \( x_1 = 2.5 \), \( f'(x_1) = -3.641 \), \( x_2 = 3.030 \)
- \( f'(x_2) = 0.178 \)
- \( x = 2.999 \)

**Step 5:** (Second Iteration)
- \( f(x) = 27.00 < f(x_1) \)

**Step 6:** (Second Iteration)
- \( f'(x) = -0.007 \) ➔ \( x_2 = 3.030 \) and \( x_1 = 2.999 \) for the next iteration

Method is faster than Powell’s quadratic since derivatives are employed. Powell method preferred if the derivatives are to be evaluated numerically.