

# Optimization Theory

MMC 52212 / MME 52106

1. Introduction

2. Single Variable Optimization

*by*

Dr. Shibayan Sarkar

Department of Mechanical Engg.

Indian School of Mines Dhanbad

# Optimal problem formulation

An **optimization problem** is the problem of finding the best solution out of all feasible solutions.

Need for Optimization

Choose Design Variable

Formulate **constraints**

Formulate objective function

Setup variable bounds

Choose an optimization algorithm

Obtain solution

.... cost, efficiency , safety

.... high sensitive to proper working design

.... represents functional relationship between design variable :

- **equality**....usually replaced by two inequality constraints
- **inequality** .... Practical cases

.... **Single objective**

**Multi objective**

- one primary single objective
- rest are constraints

.... **Upper bound**  
**lower bound**

**Constrained optimization problems** are **problems** for which a function is to be minimized or maximized **subject to constraints** .

The power of optimization methods to determine the best solution without actually testing all possible solutions comes through the use of mathematics and at the cost of performing iterative numerical calculations, using clearly defined logical procedures or algorithms implemented on computing machines.

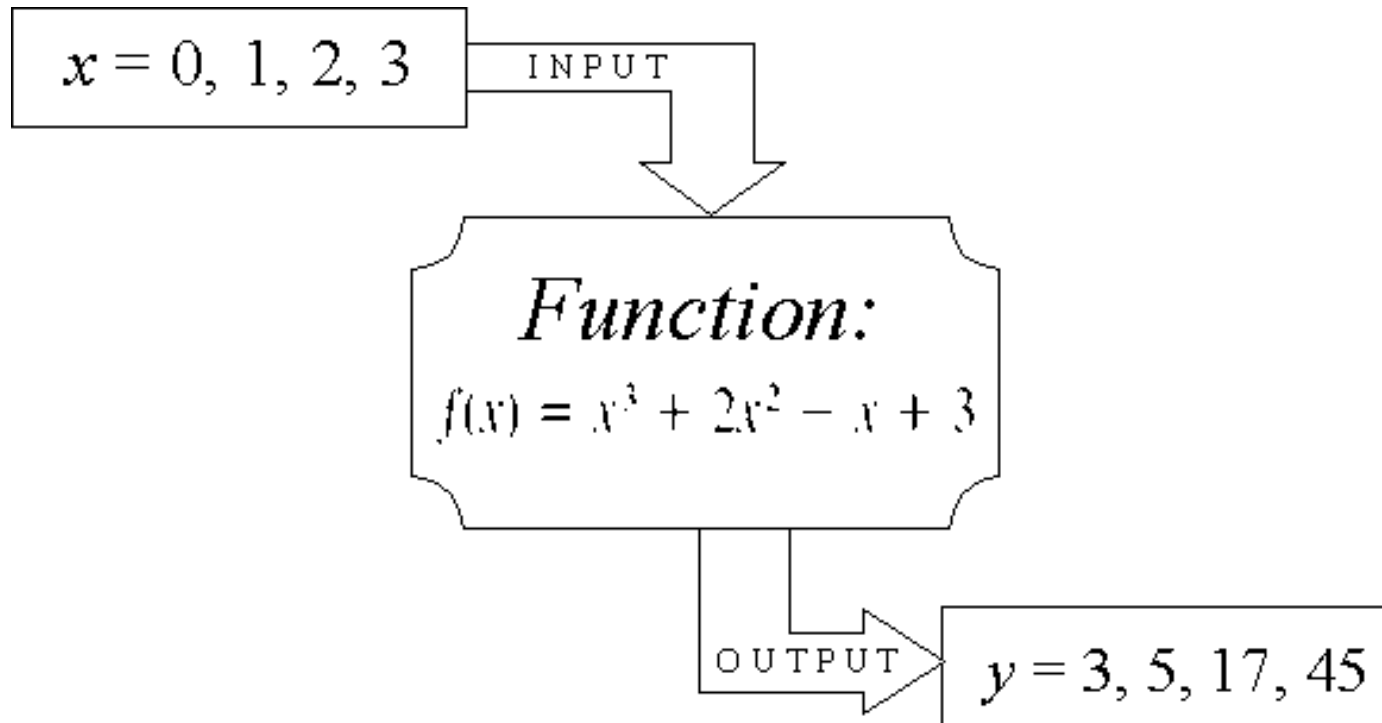
# Classification

- **Single variable optimization -**
  - *Direct method* – do not use derivative of objective function – search process
  - *Gradient based method*
- **Multivariable optimization -**
  - *unconstrained* , multivariable (Taylor series expansion) – different search methods
  - *Constrained* ... both use single variable/ multivariable repeatedly maintain search effort
    - Linear programming (objective function is linear)
    - Non- Linear programming
- **Non-traditional optimization -**
  - Genetic algorithm (GA)
  - Artificial neural network (ANN)
  - etc.

# Key features

- Following features are required to know for formulation of optimization problem (upto single variable problem ):
  - Functions
  - Optimality criteria
  - Identification of single variable optima
    - Region elimination methods
    - Polynomial approximation or point-estimation technique – search method
    - Methods requiring derivatives

# Function



- Is a rule that assigns to every choice of  $x$  a unique value  $y = f(x)$ .
- **Domain** of a function is the set of all possible input values (usually  $x$ ), which allows the function formula to work.
- **Range** is the set of all possible output values (usually  $y$ ), which result from using the function formula.

# Function

- Unconstrained and constrained function
  - Unconstrained: when domain is the entire set of real numbers  $R$
  - Constrained: domain is a proper subset of  $R$
- Continuous, discontinuous and discrete

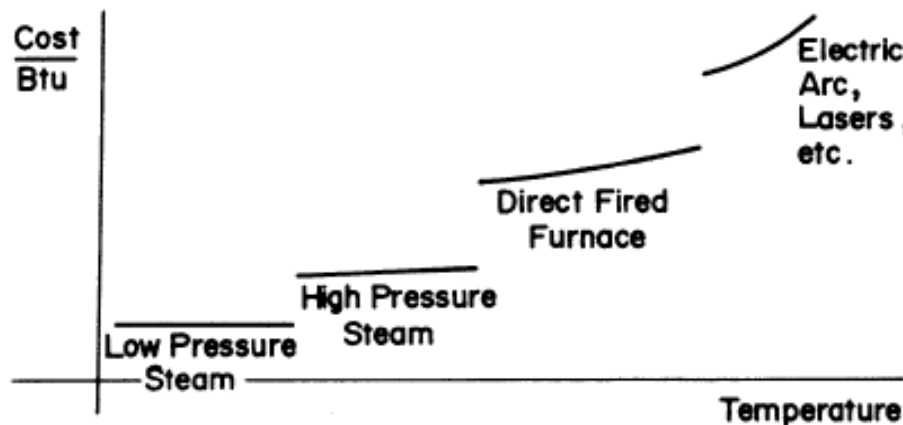


Figure 2.1. Discontinuous function.

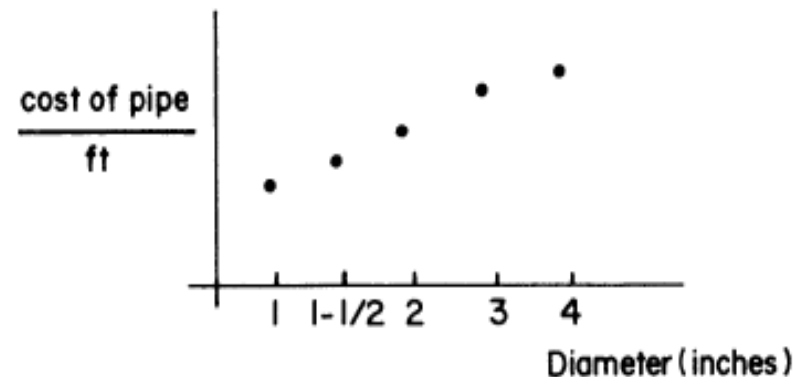


Figure 2.2. Discrete function.

# Function

- Monotonic and unimodal functions

- Monotonic:**

for any two points  $x_1$  and  $x_2$ , with  $x_1 \leq x_2$ :

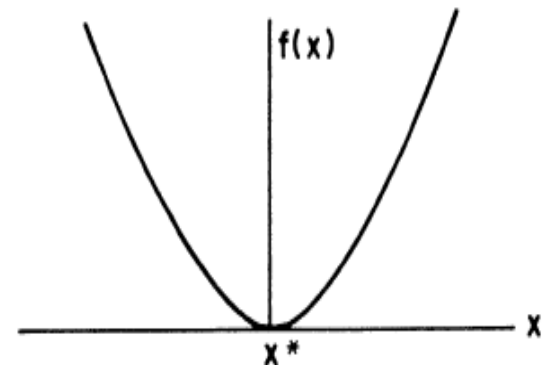
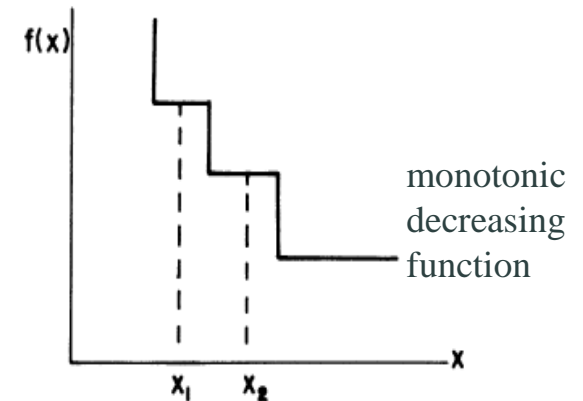
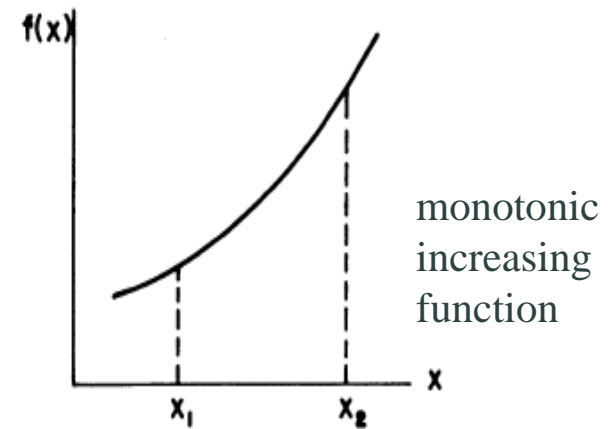
$$f(x_1) \leq f(x_2) \quad (\text{monotonically increasing})$$

$$f(x_1) \geq f(x_2) \quad (\text{monotonically decreasing})$$

- Unimodal:**

$f(x)$  is *unimodal* on the interval  $a \leq x \leq b$  if and only if it is monotonic on either side of the single optimal point  $x^*$  in the interval.

**Unimodality** is an extremely important functional property used in optimization.



# Optimality Criteria

In considering optimization problems, two questions generally must be addressed:

- **Static Optimization-** How can one determine whether a given point  $x^*$  is the optimal solution? It refers to the process of minimizing or maximizing the costs/benefits of some objective function **for one instant in time** only.
- **Dynamic Question-** If  $x^*$  is not the optimal point, then how does one go about finding a solution that is optimal? **It refers to the process of minimizing or maximizing the costs/benefits of some objective function over a period of time.** Sometimes called optimal control. **Ex. Calculus of Variation, Optimal Control, Static Optimization** to solve dynamic optimization problems **etc.**

**We are mainly concern primarily with the static question**, like developing a set of optimality criteria for determining whether a given solution is optimal.



# Optimality Criteria

## Local and global optimum (here minimum)

A function  $f(x)$  defined on a set  $S$  attains its *global minimum* at a point  $x^{**} \in S$  if and only if

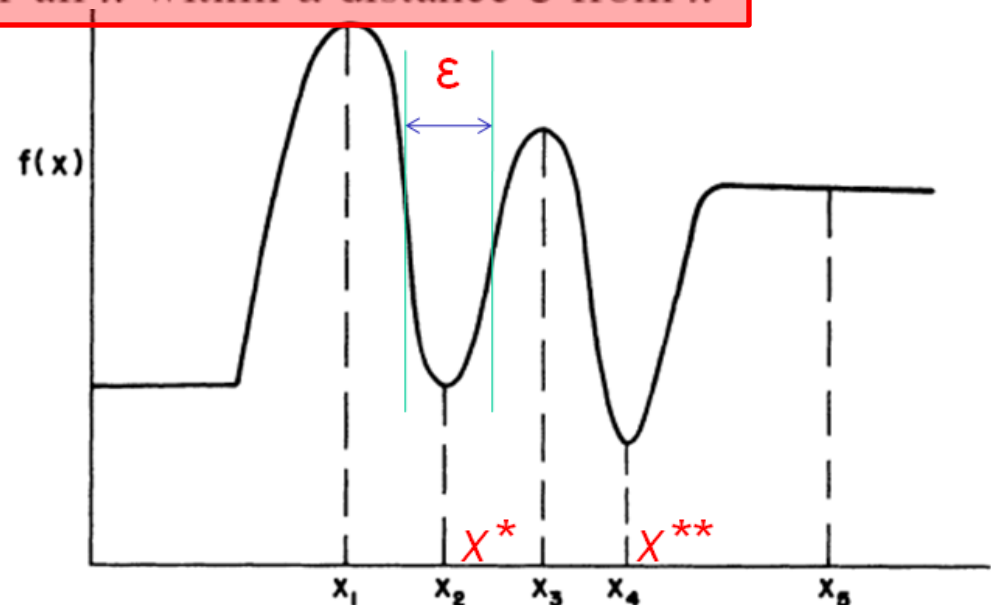
$$f(x^{**}) \leq f(x) \quad \text{for all } x \in S$$

A function  $f(x)$  defined on  $S$  has a *local minimum (relative minimum)* at a point  $x^* \in S$  if and only if

$$f(x^*) \leq f(x)$$

for all  $x$  within a distance  $\varepsilon$  from  $x^*$

that is, there exists an  $\varepsilon > 0$   
such that, for all  $x$  satisfying  
 $|x - x^*| < \varepsilon$ ,  $f(x^*) \leq f(x)$



# Identification of Single-Variable Optima

- For finding local minima (maxima)

$$\left. \frac{df}{dx} \right|_{x=x^*} = 0 \quad \text{AND} \quad \left. \frac{d^2f}{dx^2} \right|_{x=x^*} \geq 0 \ (\leq 0)$$

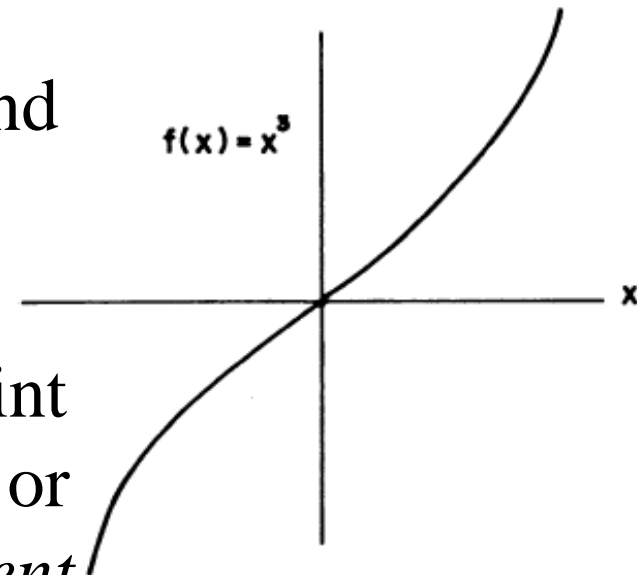
- Proof follows...
- These are necessary conditions, i.e., if they are not satisfied,  $x^*$  is not a local minimum (maximum).
- If they are satisfied, we still have no guarantee that  $x^*$  is a local minimum (maximum).

# Stationary Point and Inflection Point

- A stationary point is a point  $x^*$  at which

$$\left. \frac{df}{dx} \right|_{x=x^*} = 0$$

- An *inflection point* or *saddle-point* is a stationary point that does not correspond to a local optimum (minimum or maximum).
- To distinguish whether a stationary point is a local minimum, a local maximum, or an inflection point, we need the *sufficient* conditions of optimality.



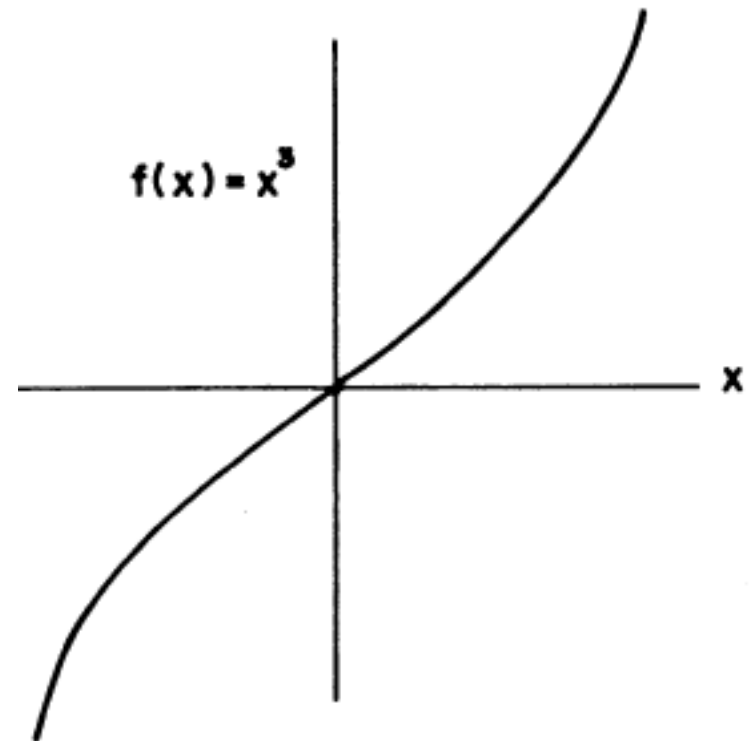
## *Theorem*

- Suppose at a point  $x^*$  the first derivative is zero and the first nonzero higher order derivative is denoted by  $n$ .
  - If  $n$  is odd, then  $x^*$  is a point of inflection.
  - If  $n$  is even, then  $x^*$  is a local optimum.
    - If that derivative is positive, then the point  $x^*$  is a local minimum.
    - If that derivative is negative, then the point  $x^*$  is a local maximum.

## Example : 1

$$f(x) = x^3$$

$$\left. \frac{df}{dx} \right|_{x=0} = 0 \quad \left. \frac{d^2f}{dx^2} \right|_{x=0} = 0 \quad \left. \frac{d^3f}{dx^3} \right|_{x=0} = 6$$



- Thus the first non-vanishing derivative is 3 (odd), and  $x = 0$  is an inflection point.

## Example : 2

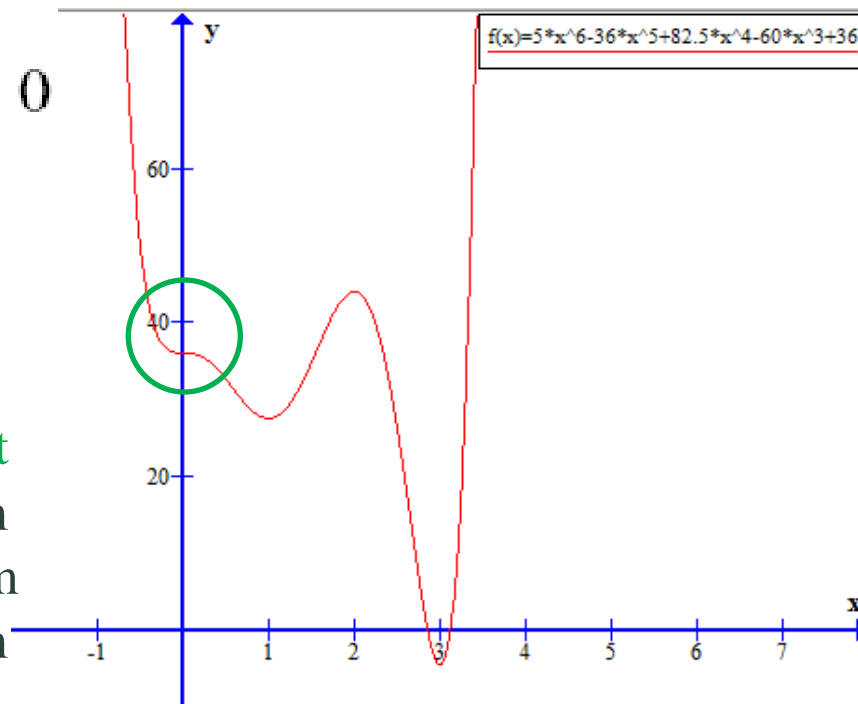
$$f(x) = 5x^6 - 36x^5 + \frac{165}{2}x^4 - 60x^3 + 36$$

$$\frac{df}{dx} = 30x^5 - 180x^4 + 330x^3 - 180x^2 = 30x^2(x-1)(x-2)(x-3)$$

Stationary points  $x = 0, 1, 2, 3$   $\left. \frac{df}{dx} \right|_{x=x^*} = 0$

$$\frac{d^2f}{dx^2} = 150x^4 - 720x^3 + 990x^2 - 360x$$

$x$	$f(x)$	$d^2f/dx^2$	$d^3f/dx^3$	
0	36	0	-360	- Inflection point
1	27.5	60	60	-Local minimum
2	44	-120	-240	-Local maximum
3	5.5	540	2340	-Local minimum



# Algorithm for finding Global Optima

Maximize  $f(x)$

Subject to  $a \leq x \leq b$

**Step 1.** Set  $df/dx = 0$  and compute all stationary points.

**Step 2.** Select all stationary points that belong to the interval  $[a, b]$ . Call them  $x_1, x_2, \dots, x_N$ . These points, along with  $a$  and  $b$ , are the only points that can qualify for a local optimum.

**Step 3.** Find the largest value of  $f(x)$  out of  $f(a), f(b), f(x_1), \dots, f(x_N)$ . This value becomes the global maximum point.

## Example : 1

Maximize  $f(x) = -x^3 + 3x^2 + 9x + 10$  in the interval  $-2 \leq x \leq 4$

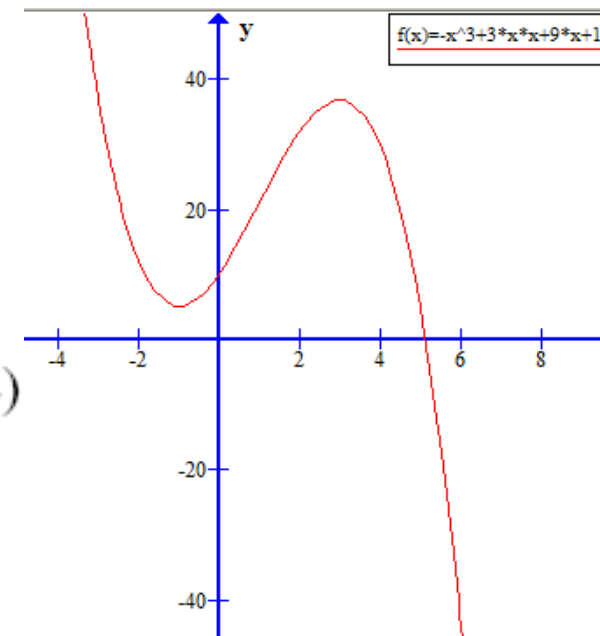
$$\frac{df}{dx} = -3x^2 + 6x + 9 = 0$$

**Stationary points  $x = -1, 3$**

To find the global maximum, evaluate  $f(x)$  at  $x = 3, -1, -2$ , and  $4$ :

$$\begin{array}{ll} f(3) = 37 & f(-1) = 5 \\ f(-2) = 12 & f(4) = 30 \end{array}$$

Hence  $x = 3$  maximizes  $f$  over the interval  $(-2, 4)$



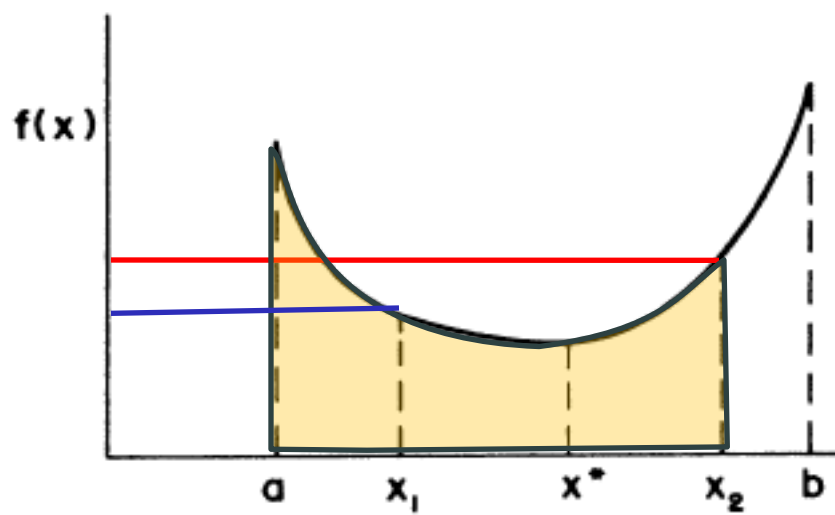
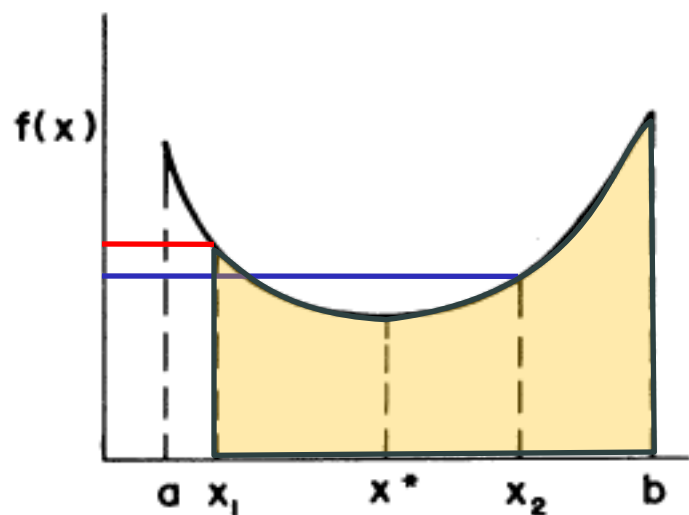


# Region Elimination Methods

Suppose  $f$  is strictly unimodal<sup>†</sup> on the interval  $a \leq x \leq b$  with a minimum at  $x^*$ . Let  $x_1$  and  $x_2$  be two points in the interval such that  $a < x_1 < x_2 < b$ .

Comparing the functional values at  $x_1$  and  $x_2$ , we can conclude:

- (i) If  $f(x_1) > f(x_2)$ , then the minimum of  $f(x)$  does not lie in the interval  $(a, x_1)$ . In other words,  $x^* \in (x_1, b)$
- (ii) If  $f(x_1) < f(x_2)$ , then the minimum does not lie in the interval  $(x_2, b)$  or  $x^* \in (a, x_2)$



# Region Elimination Methods

## How to select $x_1$ and $x_2$ ?

- Bounding Phase
  - An initial coarse search that will bound or bracket the optimum
- Interval Refinement Phase
  - A finite sequence of interval reductions or refinements to reduce the initial search interval to desired accuracy

# Bounding Phase

- Swann's method

$$x_{k+1} = x_k + 2^k \Delta \quad \text{for } k = 0, 1, 2, \dots$$

– If

$$f(x_0 - |\Delta|) \geq f(x_0) \geq f(x_0 + |\Delta|) \rightarrow \Delta \text{ is positive}$$

– Else if the inequalities are reversed  $\rightarrow \Delta$  is negative

– If

$$f(x_0 - |\Delta|) \geq f(x_0) \leq f(x_0 + |\Delta|) \rightarrow \text{the minimum lies between} \\ x_0 - |\Delta| \text{ and } x_0 + |\Delta|$$

# Bounding Phase : Example 1

Consider the problem of minimizing  $f(x) = (100 - x)^2$  given the starting point  $x_0 = 30$  and a step size  $|\Delta| = 5$ .

The sign of  $\Delta$  is determined by comparing

$$\underline{f(x_0) = f(30) = 4900}$$

$$\underline{f(x_0 + |\Delta|) = f(35) = 4225}$$

$$\underline{f(x_0 - |\Delta|) = f(25) = 5625}$$

Since

$$f(x_0 - |\Delta|) \geq f(x_0) \geq f(x_0 + |\Delta|)$$

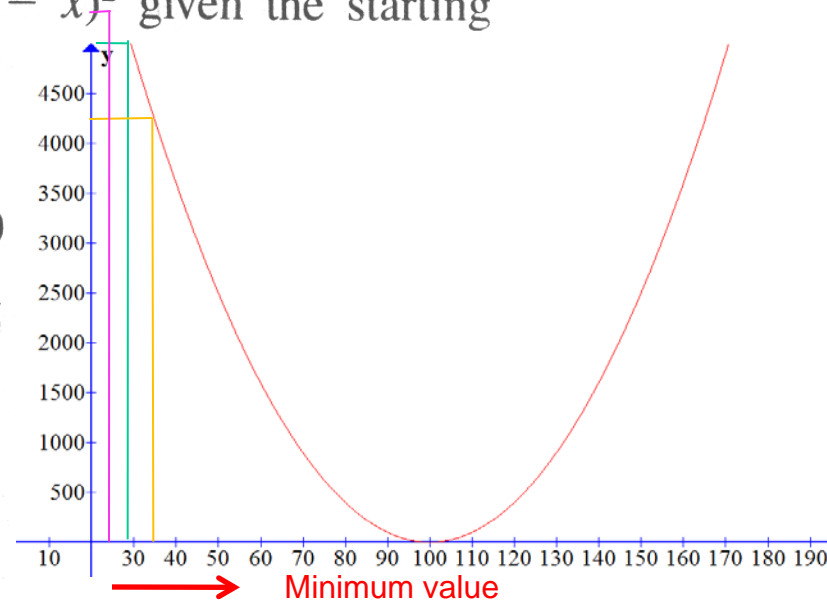
$\Delta$  must be positive, and the minimum point  $x^*$  must be greater than 30. Thus,  
 $x_1 = x_0 + \Delta = 35$ .

Next,

$$x_{k+1} = x_k + 2^k \Delta \quad \text{for } k = 0, 1, 2, \dots$$

$$x_2 = x_1 + 2\Delta = 45$$

$$f(45) = 3025 < f(x_1) \quad f(x_2) < f(x_1) \dots x^* > x_1$$



# Bounding Phase : Example 1

therefore,  $x^* > 35$ ;  $f(x_2) < f(x_1) \dots x^* > x_1$

---

$$x_3 = x_2 + 2^2\Delta = 65$$

$$f(65) = 1225 < f(x_2)$$

therefore,  $x^* > 45$ ;  $f(x_3) < f(x_2) \dots x^* > x_2$

---

$$x_4 = x_3 + 2^3\Delta = 105$$

$$f(105) = 25 < f(x_3)$$

therefore,  $x^* > 65$ ;  $f(x_4) < f(x_3) \dots x^* > x_3$

---

$$x_5 = x_4 + 2^4\Delta = 185$$

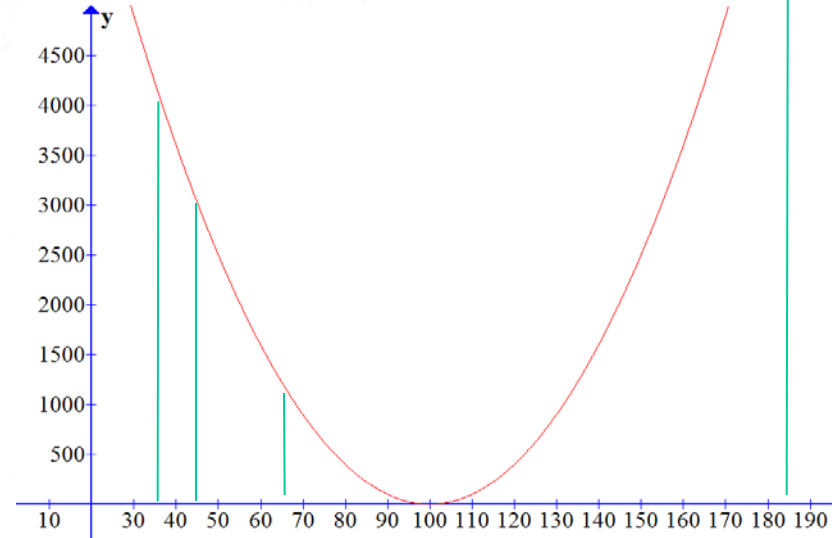
$$f(185) = 7225 > f(x_4) \quad f(x_5) > f(x_4) \dots x^* < x_4$$

therefore,  $x^* < 185$ . Consequently, in six evaluations  $x^*$  has been bracketed within the interval

---

$$65 \leq x^* \leq 185$$

$$\text{therefore} \dots x_3 < x^* < x_4$$



Note that the effectiveness of the bounding search depends directly on the step size  $\Delta$ . If  $\Delta$  is large, a poor bracket, that is, a large initial interval, is obtained. On the other hand, if  $\Delta$  is small, many evaluations may be necessary before a bound can be established.

# Interval Refinement Phase

- Interval halving

**Step 1.** Let  $x_m = \frac{1}{2}(a + b)$  and  $L = b - a$ . Compute  $f(x_m)$ .

**Step 2.** Set  $x_1 = a + \frac{1}{4}L$  and  $x_2 = b - \frac{1}{4}L$ .

Note that the points  $x_1$ ,  $x_m$ , and  $x_2$  are all equally spaced at one-fourth the interval. Compute  $f(x_1)$  and  $f(x_2)$ .

**Step 3.** Compare  $f(x_1)$  and  $f(x_m)$ .

(i) If  $f(x_1) < f(x_m)$ , then drop the interval  $(x_m, b)$  by setting  $b = x_m$ . The midpoint of the new search interval will now be  $x_1$ . Hence, set  $x_m = x_1$ . Go to step 5.

(ii) If  $f(x_1) \geq f(x_m)$ , go to step 4.

**Step 4.** Compare  $f(x_2)$  and  $f(x_m)$ .

(i) If  $f(x_2) < f(x_m)$ , drop the interval  $(a, x_m)$  by setting  $a = x_m$ . Since the midpoint of the new interval will now be  $x_2$ , set  $x_m = x_2$ . Go to step 5.

(ii) If  $f(x_2) \geq f(x_m)$ , drop the interval  $(a, x_1)$  and  $(x_2, b)$ . Set  $a = x_1$  and  $b = x_2$ . Note that  $x_m$  continues to be the midpoint of the new interval. Go to step 5.

**Step 5.** Compute  $L = b - a$ . If  $|L|$  is small, terminate. Otherwise return to step 2.

# Interval Refinement Phase

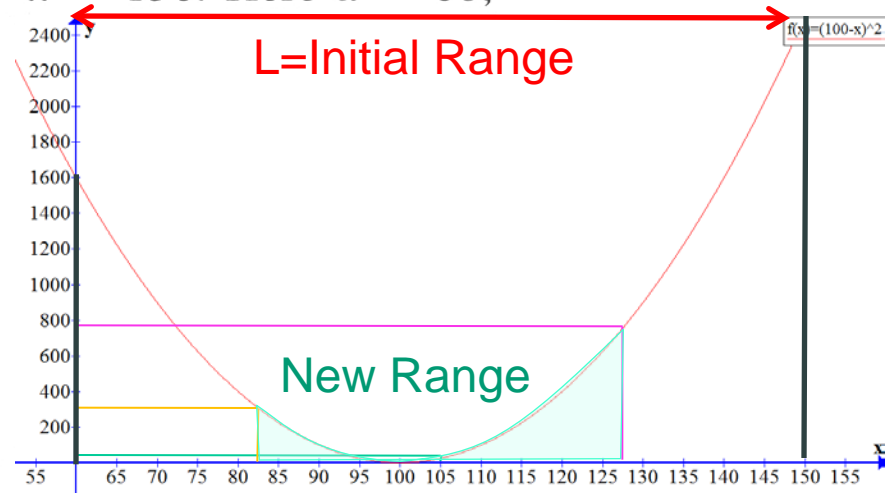
## *Remarks*

1. At each stage of the algorithm, exactly half the length of the search interval is deleted.
2. The midpoint of subsequent intervals is always equal to one of the previous trial points  $x_1$ ,  $x_2$ , or  $x_m$ . Hence, at most two functional evaluations are necessary at each subsequent step.
3. After  $n$  functional evaluations, the initial search interval will be reduced to  $(\frac{1}{2})^{n/2}$ .
4. It has been shown by Kiefer [2] that out of all equal-interval searches (two-point, three-point, four-point, etc.), the three-point search or interval halving is the most efficient.

# Interval Refinement Phase : Example 1

Minimize  $f(x) = (100 - x)^2$  over the interval  $60 \leq x \leq 150$ . Here  $a = 60$ ,  $b = 150$ , and  $L = 150 - 60 = 90$ .

$$x_m = \frac{1}{2}(60 + 150) = 105$$



**Stage 1**

$$x_1 = a + \frac{1}{4}L = 60 + \frac{90}{4} = 82.5$$

$$x_2 = b - \frac{1}{4}L = 150 - \frac{90}{4} = 127.5$$

$$f(82.5) = 306.25 > f(105) = 25$$

$$f(127.5) = 756.25 > f(105)$$

Hence, drop the intervals  $(60, 82.5)$  and  $(127.5, 150)$ . The length of the search interval is reduced from 90 to 45.



# Interval Refinement Phase : Example 1

## Stage 2

$$a = 82.5 \quad b = 127.5 \quad x_m = 105$$

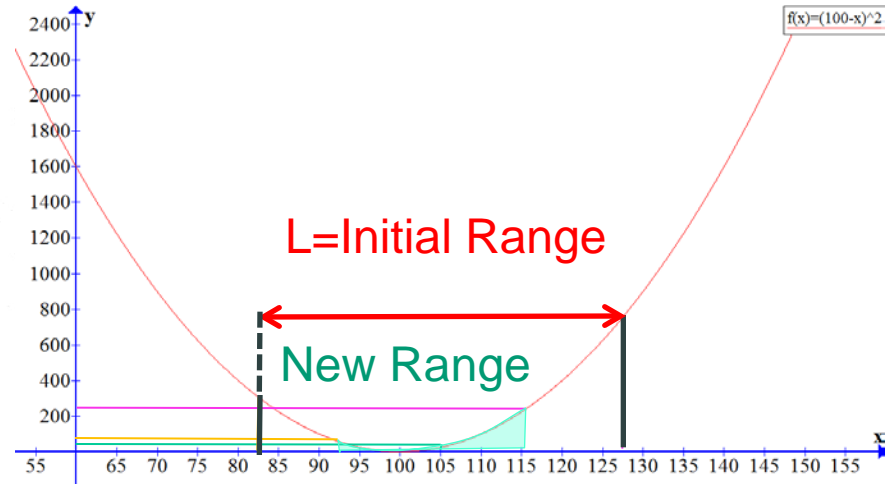
$$L = 127.5 - 82.5 = 45$$

$$x_1 = 82.5 + \frac{45}{4} = 93.75$$

$$x_2 = 127.5 - \frac{45}{4} = 116.25$$

$$f(93.75) = 39.06 > f(105) = 25$$

$$f(116.25) = 264.06 > f(105)$$



Hence, the interval of uncertainty is (93.75, 116.25).

## Stage 3

$$a = 93.75 \quad b = 116.25 \quad x_m = 105$$

$$L = 116.25 - 93.75 = 22.5$$

$$x_1 = 99.375$$

$$x_2 = 110.625$$

$$f(x_1) = 0.39 < f(105) = 25$$

Hence, delete the interval (105, 116.25). The new interval of uncertainty is now (93.75, 105), and its midpoint is 99.375 (old  $x_1$ ). Thus, in three stages (six functional evaluations), the initial search interval of length 90 has been reduced exactly to  $(90)(\frac{1}{2})^3 = 11.25$ .

# Polynomial Approximation or Point-Estimation

## Technique: by Weierstress

Using higher order polynomial .....

- Quadratic Approximation Method

If the function is smooth (unimodal and continuous), can be approximated by a polynomial and this approximating polynomial can be used to predict the location of the optimum.

Given three consecutive points  $x_1, x_2, x_3$  and their corresponding function values  $f_1, f_2, f_3$ , we seek to determine three constants  $a_0, a_1$ , and  $a_2$  such that the quadratic function

$$q(x) = a_0 + a_1(x - x_1) + a_2(x - x_1)(x - x_2)$$

agrees with  $f(x)$  at these three points. We proceed by evaluating  $q(x)$  at each of the three given points. First of all, since

$$f_1 = f(x_1) = q(x_1) = a_0$$

we have

$$a_0 = f_1$$

# Polynomial Approximation or Point-Estimation Technique: by Weierstress

Next, since

$$f_2 = f(x_2) = q(x_2) = f_1 + a_1(x_2 - x_1)$$

we have

$$a_1 = \frac{f_2 - f_1}{x_2 - x_1}$$

Finally, at  $x = x_3$ ,

$$f_3 = f(x_3) = q(x_3) = f_1 + \frac{f_2 - f_1}{x_2 - x_1} (x_3 - x_1) + a_2(x_3 - x_1)(x_3 - x_2)$$

Solving for  $a_2$ , we obtain

$$a_2 = \frac{1}{x_3 - x_2} \left( \frac{f_3 - f_1}{x_3 - x_1} - \frac{f_2 - f_1}{x_2 - x_1} \right)$$

# Polynomial Approximation or Point-Estimation Technique: by Weierstress

In the case of our quadratic approximating function,

$$\frac{dq}{dx} = a_1 + a_2(x - x_2) + a_2(x - x_1) = 0$$

can be solved to yield the estimate

$$\bar{x} = \frac{x_2 + x_1}{2} - \frac{a_1}{2a_2}$$

# Polynomial Approximation or Point-Estimation

## Technique : Example 1

Consider the estimation of the minimum of

$$f(x) = 2x^2 + \frac{16}{x}$$

on the interval  $1 \leq x \leq 5$ . Let  $x_1 = 1$ ,  $x_3 = 5$ , and choose as  $x_2$  the midpoint  $x_2 = 3$ . Evaluating the function, we obtain

$$f_1 = 18 \quad f_2 = 23.33 \quad f_3 = 53.2$$

To calculate the estimate  $\bar{x}$ , the constants  $a_1$  and  $a_2$  of the approximating function must be evaluated. Thus,

$$a_1 = \frac{23.33 - 18}{3 - 1} = \frac{8}{3}$$

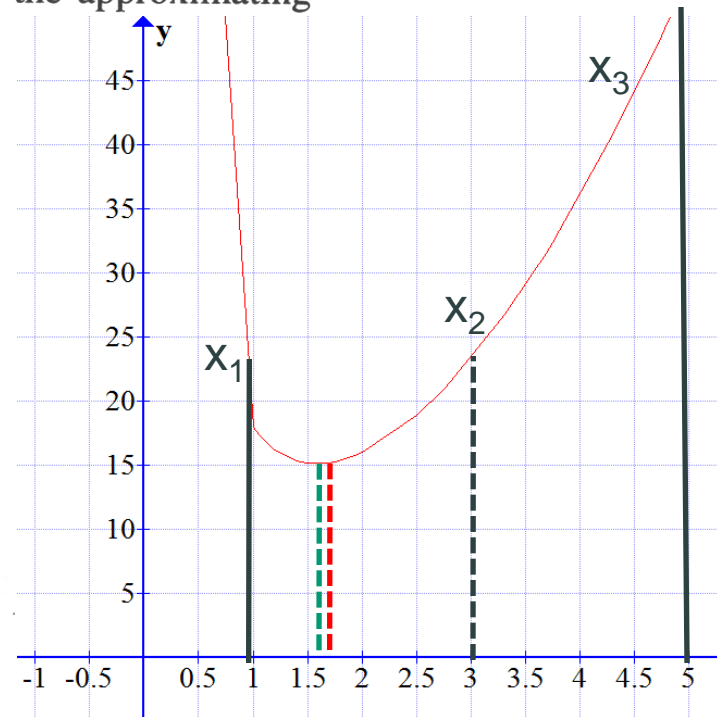
$$a_2 = \frac{1}{5 - 3} \left( \frac{53.2 - 18}{5 - 1} - \frac{8}{3} \right) = \frac{46}{15}$$

Substituting into the expression for  $\bar{x}$ ,

**Estimated minimum**

$$\bar{x} = \frac{3 + 1}{2} - \frac{8/3}{2(46/15)} = 1.565$$

The exact minimum is  $x^* = 1.5874$ .



# Successive Quadratic Estimation Method: by Powell

Reducing the interval over which the polynomial is to be approximated .....

**Step 1.** Compute  $x_2 = x_1 + \Delta x$ .

**Step 2.** Evaluate  $f(x_1)$  and  $f(x_2)$ .

**Step 3.** If  $f(x_1) > f(x_2)$ , let  $x_3 = x_1 + 2\Delta x$ .  
If  $f(x_1) \leq f(x_2)$ , let  $x_3 = x_1 - \Delta x$ .

**Step 4.** Evaluate  $f(x_3)$  and determine

Steps to find out the interval...

$$F_{\min} = \min\{f_1, f_2, f_3\}$$

$$X_{\min} = \text{point } x_i \text{ corresponding to } F_{\min}$$

**Step 5.** Use the points  $x_1, x_2, x_3$  to calculate  $\bar{x}$  using the quadratic estimation formula.

**Step 6.** Check for termination.

(a) Is  $F_{\min} - f(\bar{x})$  small enough?

(b) Is  $X_{\min} - \bar{x}$  small enough?

If both are satisfied, terminate. Otherwise, go to step 7.

**Step 7.** Save the currently best point ( $X_{\min}$  or  $\bar{x}$ ), the two points bracketing it, or the two closest to it. Relabel them and go to step 4.

Among current  $x_1$ ,  
 $x_2, x_3$

# Successive Quadratic Estimation Method :

## Example 1

$$\text{Minimize } f(x) = 2x^2 + \frac{16}{x}$$

with initial point  $x_1 = 1$  and step size  $\Delta x = 1$ . For convergence parameters use

$$\left| \frac{\text{Difference in } x}{x} \right| \leq 3 \times 10^{-2} \quad \left| \frac{\text{Difference in } F}{F} \right| \leq 3 \times 10^{-3}$$

### Iteration 1

**Step 1.**  $x_2 = x_1 + \Delta x = 2$

**Step 2.**  $f(x_1) = 18$        $f(x_2) = 16$

**Step 3.**  $f(x_1) > f(x_2)$ ; therefore set  $x_3 = 1 + 2 = 3$ .

**Step 4.**  $f(x_3) = 23.33$

$$F_{\min} = 16$$

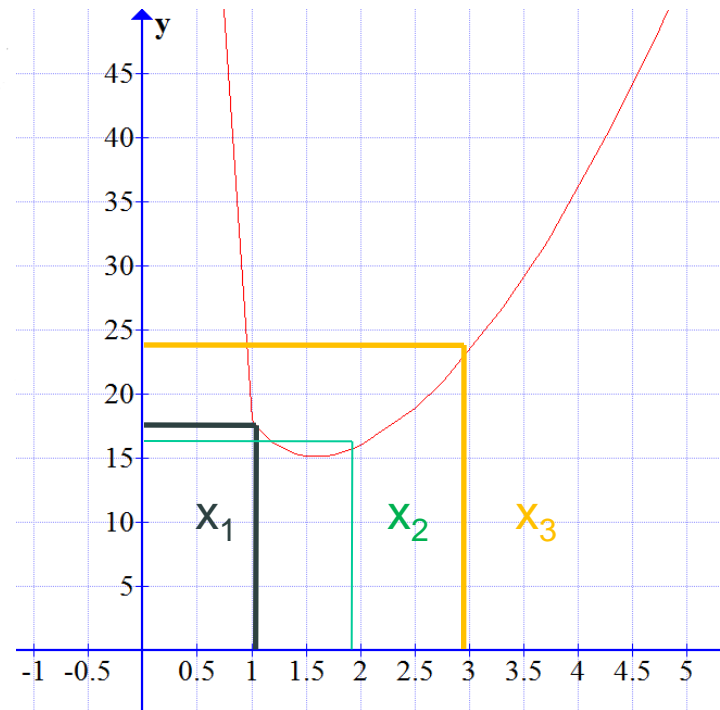
$$X_{\min} = x_2$$

**Step 5.**  $a_1 = \frac{16 - 18}{2 - 1} = -2$

$$a_2 = \frac{1}{3 - 2} \left( \frac{23.33 - 18}{3 - 1} - a_1 \right) = \frac{5.33}{2} + 2 = 4.665$$

$$\bar{x} = \frac{1 + 2}{2} - \frac{-2}{2(4.665)} = 1.5 + \frac{1}{4.665} = 1.714$$

$$f(\bar{x}) = 15.210$$



Continue.....

# Successive Quadratic Estimation Method :

## Example 1

**Step 6.** Test for termination:

$$\left| \frac{16 - 15.210}{15.210} \right| = 0.0519 > 0.003$$

Therefore continue.

Among current  $x_1, x_2, x_3$

**Step 7.** Save  $\bar{x}$ , the currently best point, and  $x_1$  and  $x_2$ , the two points that bound it. Relabel the points in order and go to iteration 2, starting with step 4.

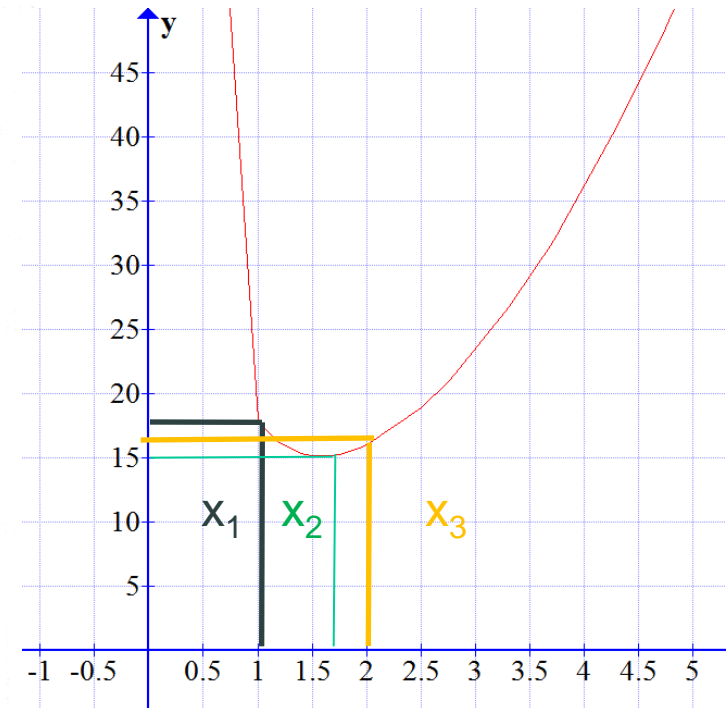
**Iteration 2**

**Step 4.**  $x_1 = 1$        $f_1 = 18$   
 $x_2 = 1.714$        $f_2 = 15.210 = F_{\min}$       and       $X_{\min} = x_2$   
 $x_3 = 2$        $f_3 = 16$

**Step 5.**  $a_1 = \frac{15.210 - 18}{1.714 - 1} = -3.908$   
 $a_2 = \frac{1}{2 - 1.714} \left( \frac{16 - 18}{2 - 1} - (-3.908) \right) = \frac{1.908}{0.286} = 6.671$   
 $\bar{x} = \frac{2.714}{2} - \frac{-3.908}{2(6.671)} = 1.357 + 0.293 = 1.650$   
 $f(\bar{x}) = 15.142$

**Step 6.** Test for termination:

$$\left| \frac{15.210 - 15.142}{15.142} \right| = 0.0045 > 0.003 \quad \text{not satisfied}$$



Continue.....



# Successive Quadratic Estimation Method :

## Example 1

**Step 7.** Save  $\bar{x}$ , the currently best point, and  $x_1 = 1$  and  $x_2 = 1.714$ , the two points that bracket it.

Among current  $x_1, x_2, x_3$

### Iteration 3

**Step 4.**  $x_1 = 1$        $f_1 = 18$   
 $x_2 = 1.65$        $f_2 = 15.142 = F_{\min}$       and       $X_{\min} = x_2$   
 $x_3 = 1.714$        $f_3 = 15.210$

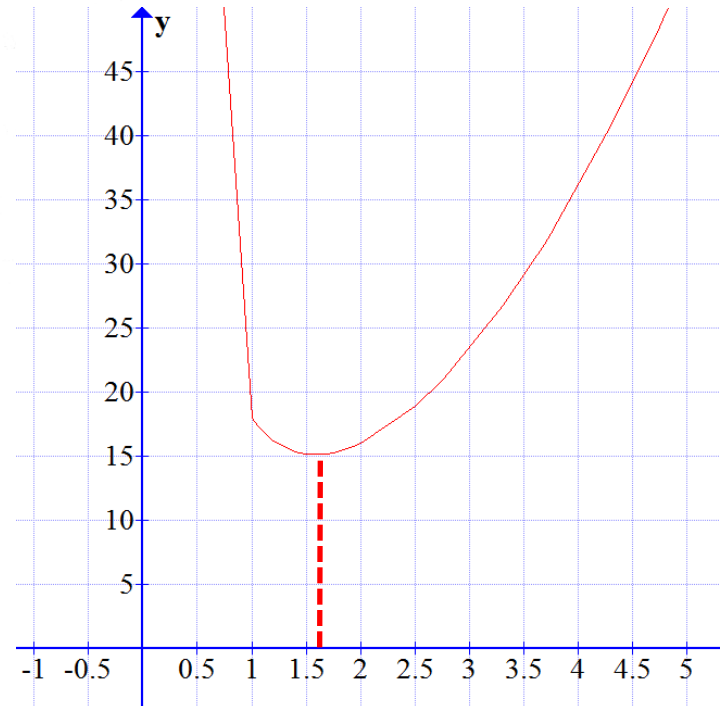
**Step 5.**  $a_1 = \frac{15.142 - 18}{1.65 - 1} = -4.397$   
 $a_2 = \frac{1}{1.714 - 1.650} \left( \frac{15.210 - 18}{1.714 - 1} - (-4.397) \right) = 7.647$   
 $\bar{x} = \frac{2.65}{2} - \frac{-4.397}{2(7.647)} = 1.6125$   
 $f(\bar{x}) = 15.123$

**Step 6.** Test for termination:

$$(i) \left| \frac{15.142 - 15.123}{15.123} \right| = 0.0013 < .003$$

$$(ii) \left| \frac{1.65 - 1.6125}{1.6125} \right| = 0.023 < 0.03$$

Therefore, terminate iterations.



# Gradient Based Method: Methods Require Derivatives - Bisection Method

Determine two points  $L$  and  $R$  such that  $f'(L) < 0$  and  $f'(R) > 0$ . The stationary point is between the points  $L$  and  $R$ . We determine the derivative of the function at the midpoint,

$$z = \frac{L + R}{2}$$

If  $f'(z) > 0$ , then the interval  $(z, R)$  can be eliminated from the search. On the other hand, if  $f'(z) < 0$ , then the interval  $(L, z)$  can be eliminated. We shall now state the formal steps of the algorithm.

Given a bounded interval  $a \leq x \leq b$  and a termination criterion  $\varepsilon$ :

**Step 1.** Set  $R = b$ ,  $L = a$ ; assume  $f'(a) < 0$  and  $f'(b) > 0$ .

**Step 2.** Calculate  $z = (R + L)/2$ , and evaluate  $f'(z)$ .

**Step 3.** If  $|f'(z)| \leq \varepsilon$ , terminate. Otherwise, if  $f'(z) < 0$ , set  $L = z$  and go to step 2. If  $f'(z) > 0$ , set  $R = z$  and go to step 2.

# Bisection Method : Example 1

Consider the problem

$$\text{Minimize } f(x) = 2x^2 + \frac{16}{x}$$

Suppose we use the **Bisection** method to determine a stationary point of  $f(x)$  starting at the point  $x_1 = 1$ :

$$f'(x) = 4x - \frac{16}{x^2} \quad f''(x) = 4 + \frac{32}{x^3}$$

Consider  $L=1$ ,  $R = 3$  Check for termination is :  $|f'(z)| \leq \epsilon$ , where  $\epsilon = 3 \times 10^{-2}$ .

**Iteration 1:**

Consider  $f'(1) = -12$ ,  $f'(3) = 12 - 16/9 = 10.22$

$$Z = (1+3)/2 = 2$$

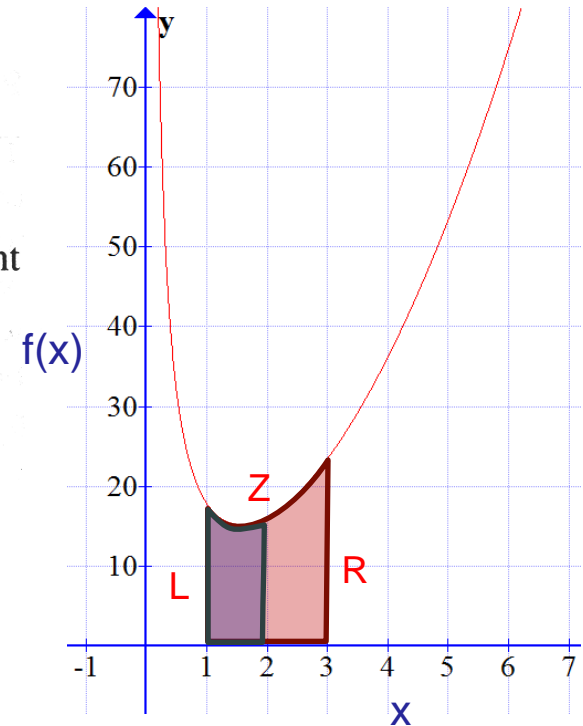
$$F'(z) = 8 - 4 = 4$$

Check for termination is :

$|f'(z)| > \epsilon$ , where  $\epsilon = 3 \times 10^{-2}$  ....so continue to next iteration.....

**Iteration 2:**

$F'(z) > 0$ ,  $R=2$ ,  $L=1$  ...proceed...



# Gradient Based Method: Newton-Raphson

- Newton-Raphson Method

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$$

The Newton-Raphson method requires the function to be twice differentiable at least.

It starts with a point  $x_1$  that is the initial estimate or approximation to the stationary point or root of the equation  $f'(x)=0$ .

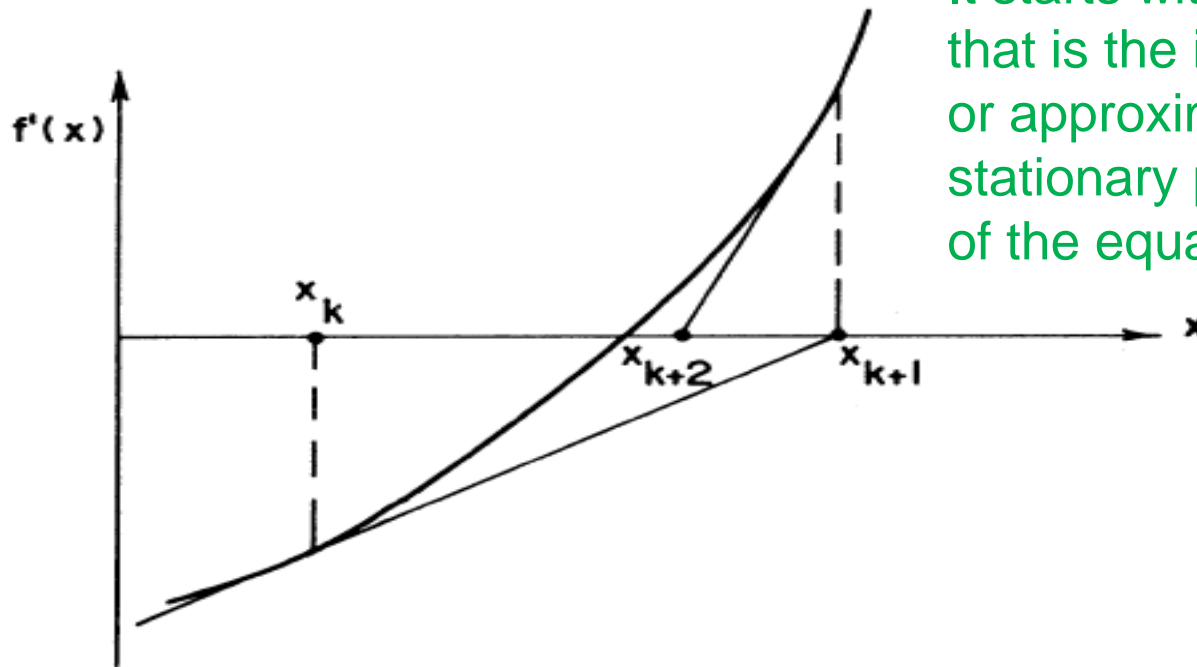


Figure 2.14. Newton-Raphson method (convergence).

# Gradient Based Method: Newton-Raphson Method : Example 1

Consider the problem

$$\text{Minimize } f(x) = 2x^2 + \frac{16}{x}$$

Suppose we use the Newton-Raphson method to determine a stationary point of  $f(x)$  starting at the point  $x_1 = 1$ :

$$f'(x) = 4x - \frac{16}{x^2} \quad f''(x) = 4 + \frac{32}{x^3}$$

**Step 1.**  $x_1 = 1$      $f'(x_1) = -12$      $f''(x_1) = 36$

$$x_2 = 1 - \frac{-12}{36} = 1.33$$

Check for termination is :  $|f'(x)| \leq \varepsilon$ .

**Step 2.**  $x_2 = 1.33$      $f'(x_2) = -3.73$      $f''(x_2) = 17.6$

$$x_3 = 1.33 - \frac{-3.73}{17.6} = 1.54$$

We continue until  $|f'(x_k)| < \varepsilon$ , where  $\varepsilon$  is a prespecified tolerance.

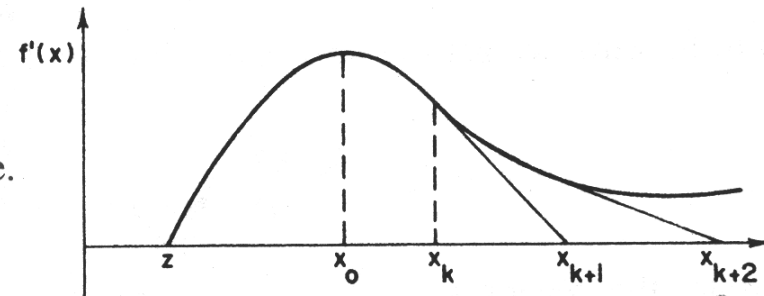
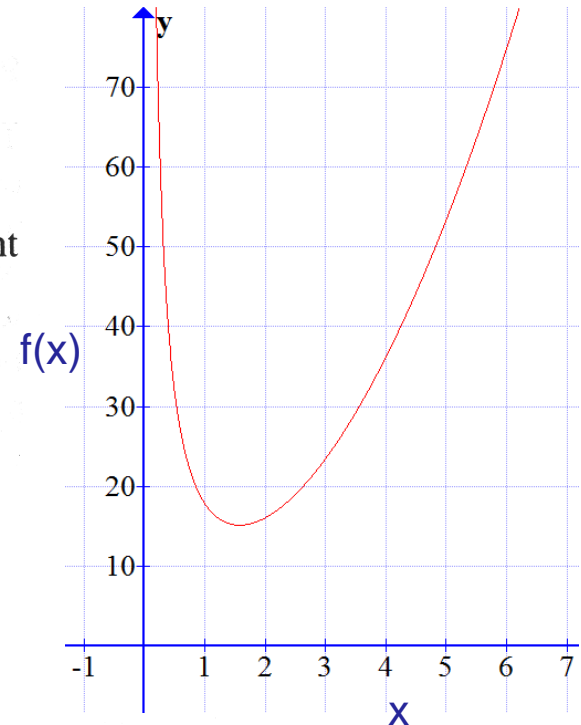


Figure 2.15. Newton-Raphson method (divergence).

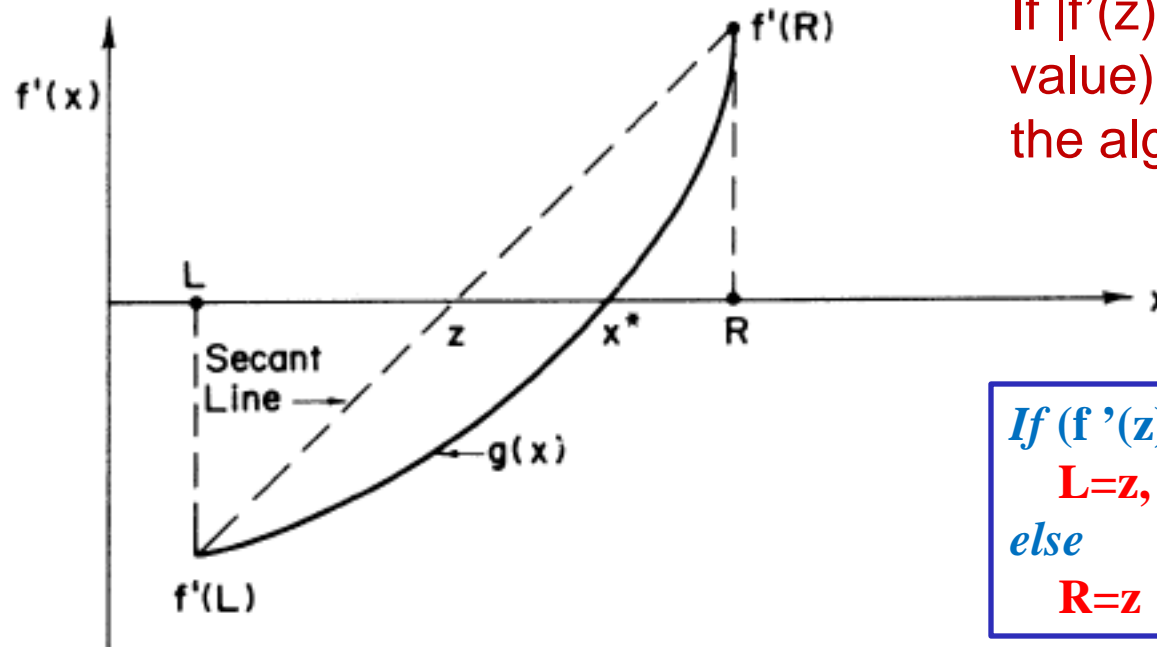
# Secant Method

Bisection method

The *secant method* combines Newton's method with a region elimination scheme for finding a root of the equation  $f'(x) = 0$  in the interval  $(a, b)$  if it exists.

The next approximation to the stationary point  $x^*$  is given as

$$z = R - \frac{f'(R)}{[f'(R) - f'(L)]/(R - L)}$$



If  $|f'(z)| \leq \epsilon$  (a given value), terminate the algorithm.

If  $f'(z) < 0$

$L = z$ ,

else

$R = z$

Figure 2.16. Secant method.

# Secant Method

Minimize  $f(x) = 2x^2 + \frac{16}{x}$  over the interval  $1 \leq x \leq 5$

Suppose,  $\epsilon = 0.3$

$$f'(x) = \frac{df(x)}{dx} = 4x - \frac{16}{x^2}$$

$\epsilon$  value or number of iteration will be given.

## Iteration 1

Step 1.  $R = 5$        $L = 1$        $f'(R) = 19.36$        $f'(L) = -12$

Step 2.  $z = 5 - \frac{19.36}{(19.36 + 12)/4} = 2.53$

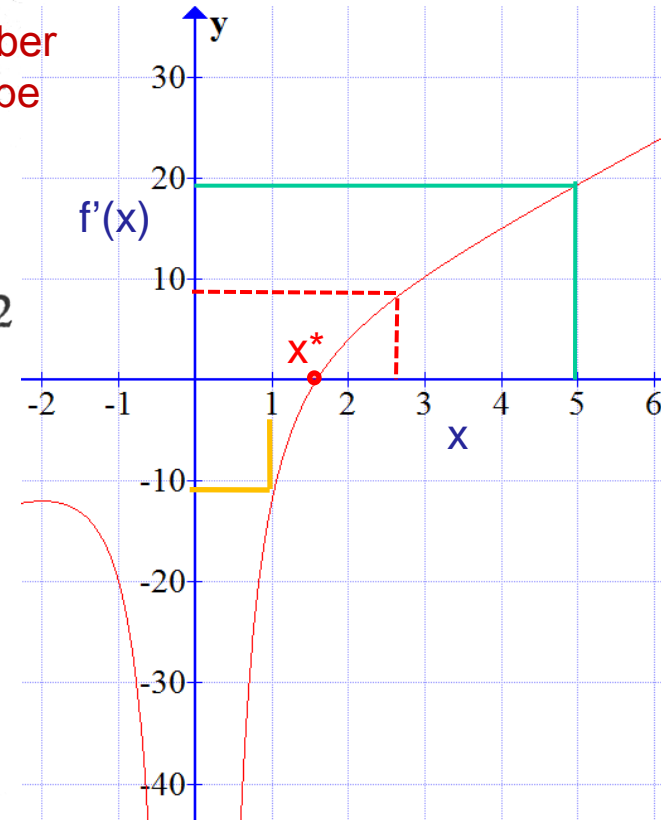
Step 3.  $f'(z) = 7.62 > 0$ ; set  $R = 2.53$ .

Iteration 2      New  $R=2.53$ ,  $f'(R)=7.62$

Step 2.  $z = 2.53 - \frac{7.62}{(7.62 + 12)/1.53} = 1.94$

Step 3.  $f'(z) = 3.51 > 0$ ; set  $R = 1.94$ .

Continue until  $|f'(z)| \leq \epsilon$ .



# Polynomial approximation method: Cubic Search Method (Third order polynomial)

Basic operation is same with quadratic method, but here both function and the derivative value will be used to find out optimal solution.

The cubic search starts with an arbitrary point  $x_1$  and finds another point  $x_2$  by a bounding search such that the derivatives  $f'(x_1)$  and  $f'(x_2)$  are of opposite sign. In other words, the stationary point  $\bar{x}$  where  $f'(x) = 0$  is bracketed between  $x_1$  and  $x_2$ . A cubic approximation function of the form

$$\bar{f}(x) = a_0 + a_1(x - x_1) + a_2(x - x_1)(x - x_2) + a_3(x - x_1)^2(x - x_2)$$

$$\frac{d\bar{f}(x)}{dx} = a_1 + a_2(x - x_1) + a_2(x - x_2) + a_3(x - x_1)^2 + 2a_3(x - x_1)(x - x_2)$$

$$f_1 = f(x_1) = a_0$$

$$f_2 = f(x_2) = a_0 + a_1(x_2 - x_1)$$

$$f'_1 = f'(x_1) = a_1 + a_2(x_1 - x_2)$$

$$f'_2 = f'(x_2) = a_1 + a_2(x_2 - x_1) + a_3(x_2 - x_1)^2$$



# Cubic Search Method

Given an initial point  $x_0$ , positive step size  $\Delta$ , and termination parameters  $\varepsilon_1$  and  $\varepsilon_2$ , the formal steps of the cubic search algorithm are as follows:

**Step 1.** Compute  $f'(x_0)$ .

If  $f'(x_0) < 0$ , compute  $x_{K+1} = x_K + 2^K \Delta$  for  $K = 0, 1, \dots$

If  $f'(x_0) > 0$ , compute  $x_{K+1} = x_K - 2^K \Delta$  for  $K = 0, 1, 2, \dots$

**Step 2.** Evaluate  $f'(x)$  for points  $x_{K+1}$  for  $K = 0, 1, 2, \dots$  until a point  $x_M$  is reached at which  $f'(x_{M-1})f'(x_M) \leq 0$ .

Then set  $x_1 = x_{M-1}$ ,  $x_2 = x_M$ .

Compute  $f_1, f_2, f'_1$ , and  $f'_2$ .

**Step 3.** Calculate the stationary point  $\bar{x}$  of the cubic approximating function using Eq. (2.10).

**Step 4.** If  $f(\bar{x}) < f(x_1)$ , go to step 5. Otherwise, set  $\bar{x}_{new} = \bar{x}_{old} + \frac{1}{2}(\bar{x}_{old} - x_1)$  until  $f(\bar{x}_{new}) \leq f(x_1)$  is achieved.

**Step 5.** Termination check:

If  $|f'(\bar{x})| \leq \varepsilon_1$  and  $|(\bar{x} - x_1)/\bar{x}| \leq \varepsilon_2$ , stop. Otherwise, set

(i)  $x_2 = x_1$  and  $x_1 = \bar{x}$  if  $f'(\bar{x})f'(x_1) < 0$

(ii)  $x_1 = \bar{x}$  if  $f'(\bar{x})f'(x_2) < 0$

In either case, continue with step 3.

# Cubic Search Method : Example 1

**Q.**

$$\text{Minimize } f(x) = 2x^2 + \frac{16}{x}$$

with the initial point  $x_0 = 1$  and step size  $\Delta = 1$ .

For convergence parameter use

$$\varepsilon_1 = 10^{-2} \quad \varepsilon_2 = 3 \times 10^{-2}$$

$$f'(x) = \frac{df}{dx} = 4x - \frac{16}{x^2}$$

## Iteration 1

**Step 1.**  $f'(1) = -12 < 0$ . Hence,  $x_1 = 1 + 1 = 2$ .

**Step 2.**  $f'(2) = 4$

Since  $f'(1)f'(2) = -48 < 0$ , a stationary point has been bracketed between 1 and 2. Set  $x_1 = 1$ ,  $x_2 = 2$ . Then,  $f_1 = 18$ ,  $f_2 = 16$ ,  $f'_1 = -12$ ,  $f'_2 = 4$ .

**Step 3.**  $z = \frac{3}{1}(18 - 16) + (-12) + 4 = -2$

$$w = +[4 - (-12)(4)]^{1/2} = (52)^{1/2} = 7.211$$

$$\mu = \frac{4 + 7.211 - (-2)}{4 - (-12) + 2(7.211)} = 0.4343$$

$$\bar{x} = 2 - 0.4343(2 - 1) = 1.5657$$

**Step 4.**  $f(1.5657) = 15.1219 < f(x_1) = 18$

Therefore, continue.

A closed form solution to the stationary point  $\bar{x}$  of the approximating cube is obtained as follows:

$$\bar{x} = \begin{cases} x_2 & \text{if } \mu < 0 \\ x_2 - \mu(x_2 - x_1) & \text{if } 0 \leq \mu \leq 1 \\ x_1 & \text{if } \mu > 1 \end{cases} \quad (2.10)$$

Where,

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

$$z = \left( \frac{3(f_1 - f_2)}{x_2 - x_1} \right) + f'_1 + f'_2$$

$$w = \begin{cases} (z^2 - f'_1 f'_2)^{1/2} & \text{if } x_1 < x_2 \\ -(z^2 - f'_1 f'_2)^{1/2} & \text{if } x_1 > x_2 \end{cases}$$

**Step 5.** Termination check:

$$f'(1.5657) = -0.2640 \quad \text{clearly not terminated}$$

Since

$$f'(\bar{x})f'(x_2) = (-0.2640)(4) < 0$$

$$\text{set } x_1 = \bar{x} = 1.5657.$$

# Cubic Search Method : Example 1

## Iteration 2

$$\text{Step 3. } z = \frac{3}{0.4343} (15.1219 - 16) + (-0.2640) + 4 = -2.3296$$

$$w = +[(2.3296)^2 - (-0.2640)(4)]^{1/2} = 2.5462$$

$$\mu = \frac{4 + 2.5462 - (-2.3296)}{4 - (-0.2640) + 2(2.5462)} = 0.9486$$

$$\bar{x} = 2 - 0.9486(2 - 1.5657) = 1.5880$$

$$\text{Step 4. } f(1.5880) = 15.1191 < f(x_1) = 15.1219$$

Therefore, continue.

Step 5. Termination check:

$$f'(1.5880) = 0.0072 < 10^{-2}$$

$$\left| \frac{1.5880 - 1.5657}{1.5880} \right| = 0.0140 < 3 \times 10^{-2}$$

Therefore, terminate.

Note that using the same two points,  $x_1 = 1$  and  $x_2 = 2$ , the quadratic search of Example 2.9 returned an estimate  $\bar{x} = 1.714$ , whereas the cubic search produced 1.5657. The exact minimum is 1.5874, indicating the clear superiority of the higher order polynomial fit.

# References

1. Optimization for Engineering Design by Kalyanmoy Deb
2. Engineering Optimization: Methods and Applications by A. Ravindran, K, M, Ragsdell, G.V. Reklaitis
3. Engineering Optimization: Theory and Practice by S. S. Rao