CH 4020: Optimization Techniques I
CH331: Engineering Elective 4 - Optimization

Spring 2022

Instructor: Dr. Kishalay Mitra

Global Optimization & Knowledge Unearthing Lab (GOKUL)

Department of Chemical Engineering

Indian Institute of Technology Hyderabad

(kishalay@iith.ac.in)

https://sites.google.com/site/kishalaymitra/

How far with analytical methods…

- **If objectives, constraints are**
  - **simple to express and**
  - **possible to express explicitly in terms of decision variables**

$$\text{Minimize } f(\mathbf{X}) = \sum_{i=1}^{11} \rho x_i l_i$$

constraints

$$g_j(\mathbf{X}) = |u_j(\mathbf{X})| - \delta \le 0, \qquad j = 1, 2, \ldots, 10$$

$$x_i \ge 0, \qquad i = 1, 2, \ldots, 11$$

$$(4x_4 + x_6 + x_7)u_1 + \sqrt{3}(x_6 - x_7)u_2 - 4x_4u_3 - x_7u_7 + \sqrt{3}x_7u_8 = 0$$

$$\sqrt{3}(x_6 - x_7)u_1 + 3(x_6 + x_7)u_2 + \sqrt{3}x_7u_7 - 3x_7u_8 = -\frac{4Rl}{E}$$

$$\begin{aligned} &-4x_4u_1 + (4x_4 + 4x_5 + x_8 + x_9)u_3 + \sqrt{3}(x_8 - x_9)u_4 - 4x_5u_5 \\ &- x_8u_7 - \sqrt{3}x_8u_8 - x_9u_9 + \sqrt{3}x_9u_{10} = 0 \end{aligned}$$

$$\begin{aligned} &\sqrt{3}(x_8 - x_9)u_3 + 3(x_8 + x_9)u_4 - \sqrt{3}x_8u_7 \\ &- 3x_8u_8 + \sqrt{3}x_9u_9 - 3x_9u_{10} = 0 \end{aligned}$$

$$\begin{aligned} &-4x_5u_3 + (4x_5 + x_{10} + x_{11})u_5 + \sqrt{3}(x_{10} - x_{11})u_6 \\ &- x_{10}u_9 - \sqrt{3}x_{10}u_{10} = \frac{4Ql}{E} \end{aligned}$$

$$\sqrt{3}(x_{10} - x_{11})u_5 + 3(x_{10} + x_{11})u_6 - \sqrt{3}x_{10}u_9 - 3x_{10}u_{10} = 0$$

$$\begin{aligned} &-x_7u_1 + \sqrt{3}x_7u_2 - x_8u_3 - \sqrt{3}x_8u_4 + (4x_1 + 4x_2 \\ &+ x_7 + x_8)u_7 - \sqrt{3}(x_7 - x_8)u_8 - 4x_2u_9 = 0 \end{aligned}$$

$$\begin{aligned} &\sqrt{3}x_7u_1 - 3x_7u_2 - \sqrt{3}x_8u_3 - 3x_8u_4 - \sqrt{3}(x_7 - x_8)u_7 \\ &+ 3(x_7 + x_8)u_8 = 0 \end{aligned}$$

$$\begin{aligned} &-x_9u_3 + \sqrt{3}x_9u_4 - x_{10}u_5 - \sqrt{3}x_{10}u_6 - 4x_2u_7 \\ &+ (4x_2 + 4x_3 + x_9 + x_{10})u_9 - \sqrt{3}(x_9 - x_{10})u_{10} = 0 \end{aligned}$$

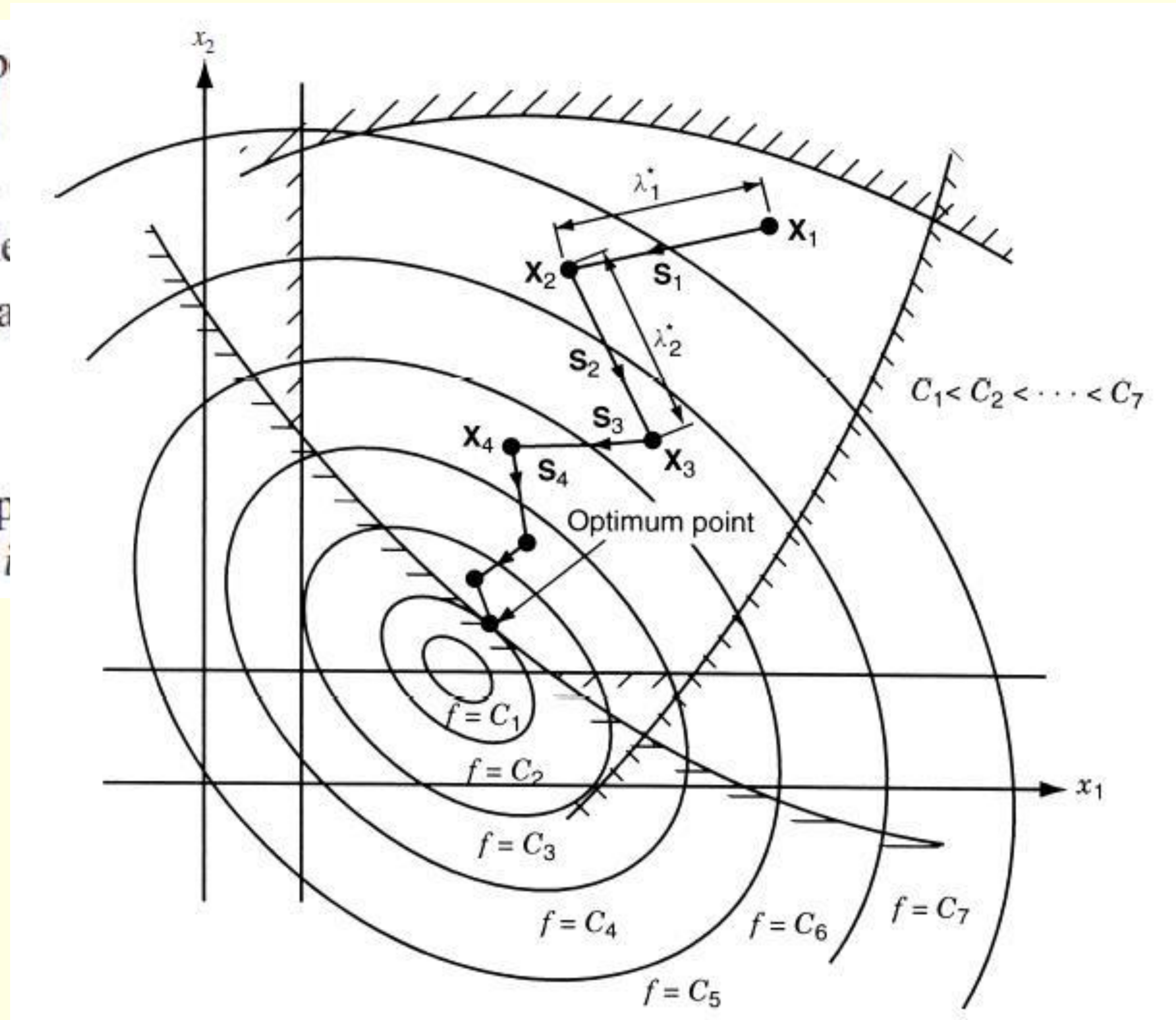
$$\begin{aligned} &\sqrt{3}x_9u_3 - 3x_9u_4 - \sqrt{3}x_{10}u_5 - 3x_{10}u_6 - \sqrt{3}(x_9 - x_{10})u_9 \\ &+ 3(x_9 + x_{10})u_{10} = -\frac{4Sl}{E} \end{aligned}$$

Numerical Optimization

- **Basic philosophy in any numerical optimization**

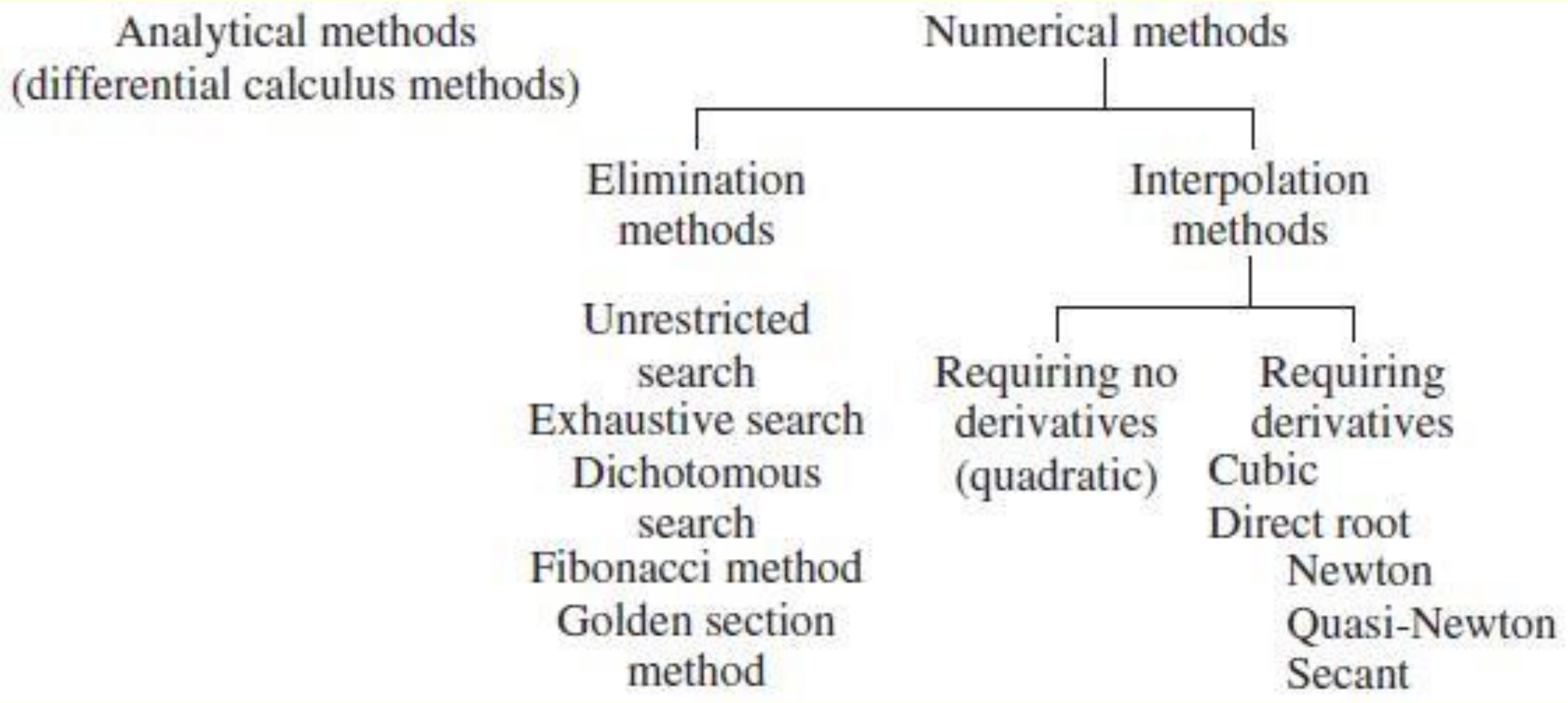
1. Start with an initial trial p
2. Find a suitable direction
   direction of the optimum.
3. Find an appropriate step le
4. Obtain the new approxima
5. Test whether X_{i+1} is op
   Otherwise, set a new $i = i$

Way of calculating search directions and step length are going to give us different optimization techniques



One Dimensional Optimization

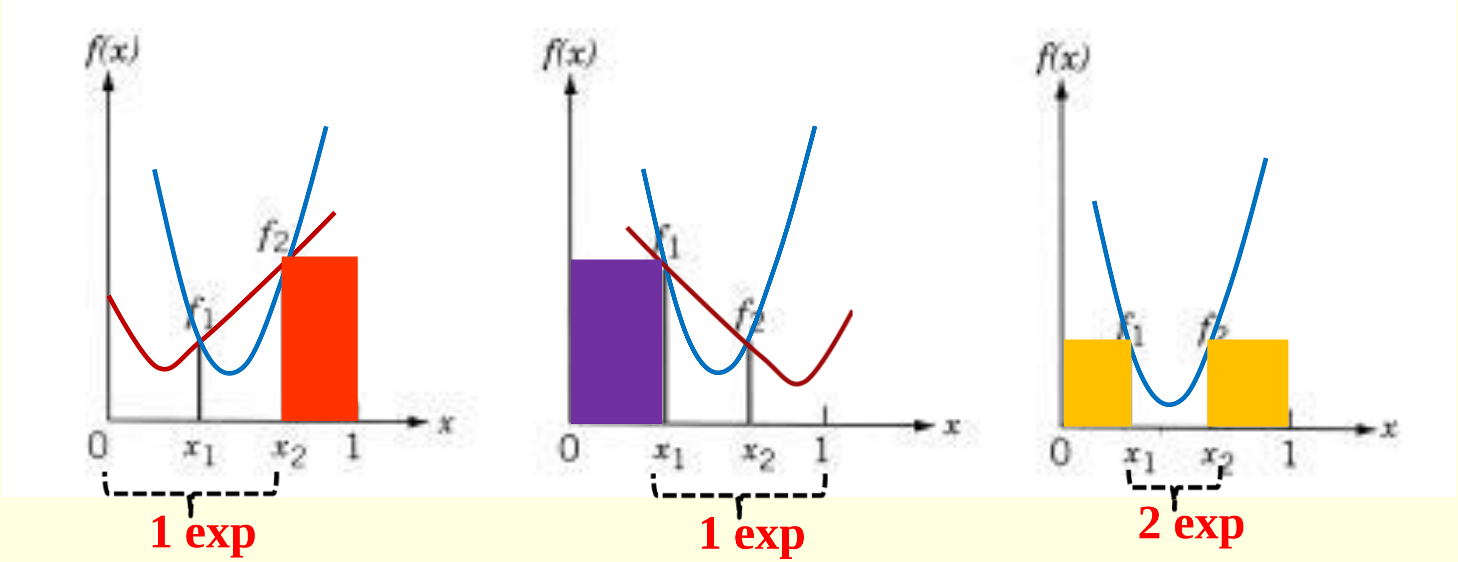
- **Classification**



Unimodality

- **Function which is having only one peak (maximization) or one valley (minimization) in a given interval**
- **Given 2 values of the variable on the same side of the optimum, one nearer to the optimum gives smaller value in case of minimization**

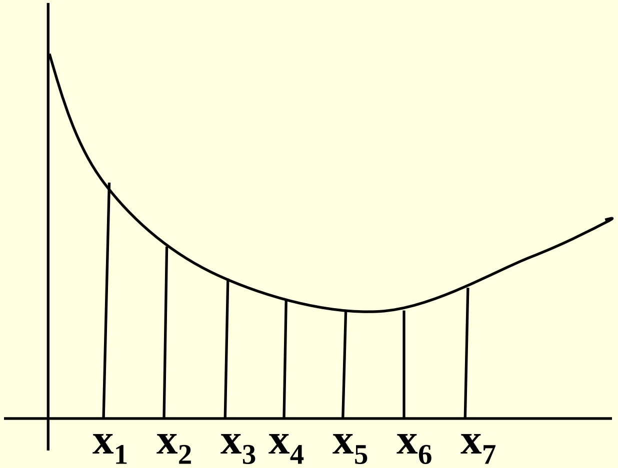
A function $f(x)$ is unimodal if (*i*) $x_1 < x_2 < x^*$ implies that $f(x_2) < f(x_1)$, and (ii) $x_2 > x_1 > x^*$ implies that $f(x_1) < f(x_2)$, where x^* is the minimum point.



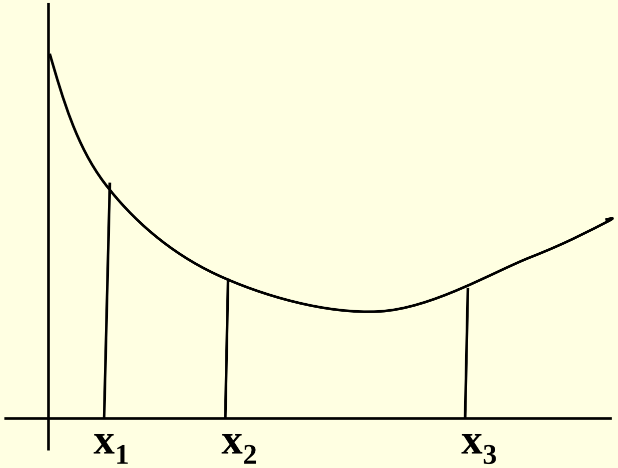
Area Elimination

Unrestricted Search (how to bracket the optimum when the ranges are not given)

Fixed Step Size



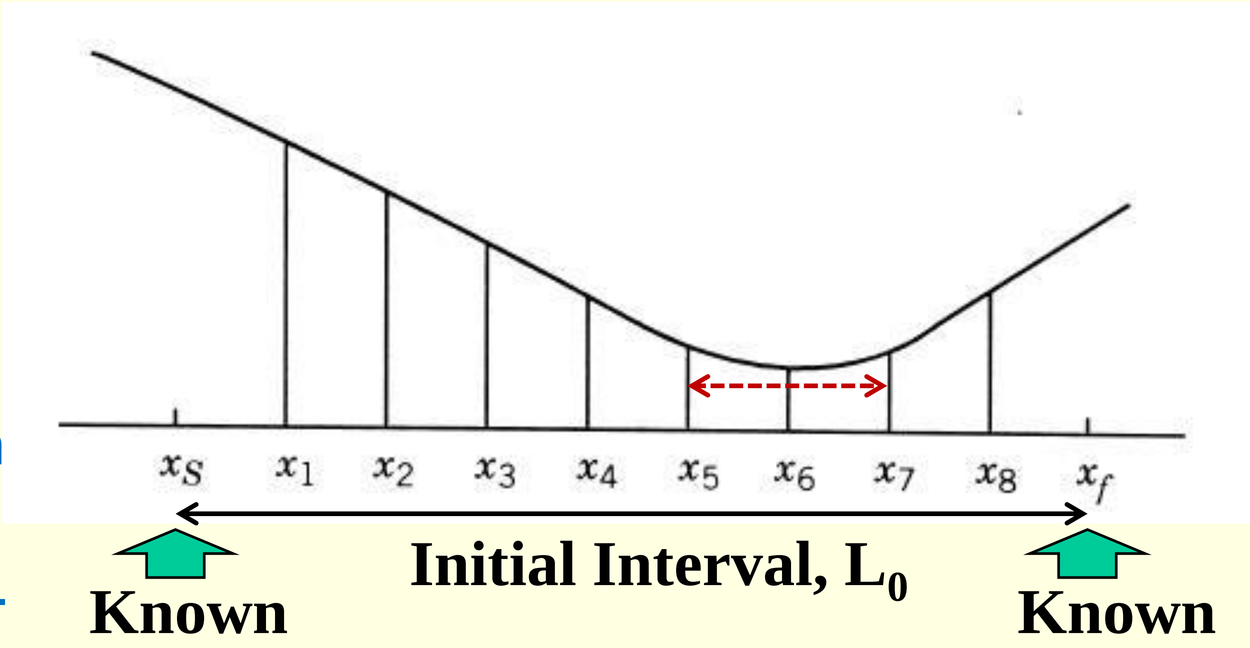
Accelerated Step Size



Area Elimination

Exhaustive Search

- Bounds on the search space is known
- Equally spaced n points are placed in the initial interval, L_0 giving rise to n+1 segments
- Final interval of uncertainty L_n [x_5, x_7] having 2 segments



$$\frac{L_n}{L_0} = \frac{2}{n+1}$$

Exhaustive Search Example

Example 5.4 Find the minimum of $f = x(x - 1.5)$ in the interval (0.0, 1.00) to within 10% of the exact value.

SOLUTION If the middle point of the final interval of uncertainty is taken as the approximate optimum point, the maximum deviation could be $1/(n+1)$ times the initial interval of uncertainty. Thus to find the optimum within 10% of the exact value, we should have

$$\frac{1}{n+1} \leq \frac{1}{10} \quad \text{or} \quad n \geq 9$$

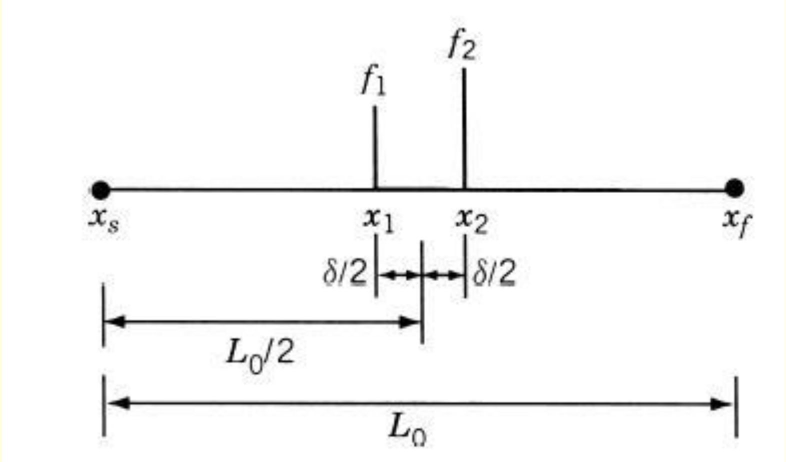
By taking $n = 9$, the following function values can be calculated:

i	1	2	3	4	5	6	7	8	9
x_i	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$f_i = f(x_i)$	−0.14	−0.26	−0.36	−0.44	−0.50	−0.54	−0.56	−0.56	−0.54

Since $x_7 = x_8$, the assumption of unimodality gives the final interval of uncertainty as $L_9 = (0.7, 0.8)$. By taking the middle point of L_9 (i.e., 0.75) as an approximation to the optimum point, we find that it is, in fact, the true optimum point.

Dichotomous Search

- Mid point of the interval identified
- 2 points x_1, x_2 created ($\delta/2$) away from the mid point
- Based on these function values, certain area [x_2, x_f] is eliminated
- Process repeated again



$$x_1 = \frac{L_0}{2} - \frac{\delta}{2}$$

$$x_2 = \frac{L_0}{2} + \frac{\delta}{2}$$

Number of experiments	2	4	6
Final interval of uncertainty	$\frac{1}{2}(L_0 + \delta)$	$\frac{1}{2}\left(\frac{L_0 + \delta}{2}\right) + \frac{\delta}{2}$	$\frac{1}{2}\left(\frac{L_0 + \delta}{4} + \frac{\delta}{2}\right) + \frac{\delta}{2}$

$$L_n = \frac{L_0}{2^{n/2}} + \delta\left(1 - \frac{1}{2^{n/2}}\right)$$

Dichotomous Search Example

Example 5.5 Find the minimum of $f = x(x - 1.5)$ in the interval (0.0, 1.00) to within 10% of the exact value.

SOLUTION The ratio of final to initial intervals of uncertainty is given by [from Eq. (5.3)]

$$\frac{L_n}{L_0} = \frac{1}{2^{n/2}} + \frac{\delta}{L_0}\left(1 - \frac{1}{2^{n/2}}\right)$$

where δ is a small quantity, say 0.001, and n is the number of experiments. If the middle point of the final interval is taken as the optimum point, the requirement can be stated as

$$\frac{1}{2}\frac{L_n}{L_0} \le \frac{1}{10}$$

i.e.,

$$\frac{1}{2^{n/2}} + \frac{\delta}{L_0}\left(1 - \frac{1}{2^{n/2}}\right) \le \frac{1}{5}$$

Since $\delta = 0.001$ and $L_0 = 1.0$, we have

$$\frac{1}{2^{n/2}} + \frac{1}{1000}\left(1 - \frac{1}{2^{n/2}}\right) \le \frac{1}{5}$$

i.e.,

$$\frac{999}{1000}\frac{1}{2^{n/2}} \le \frac{995}{5000} \quad \text{or} \quad 2^{n/2} \ge \frac{999}{199} \simeq 5.0$$

Since n has to be even, this inequality gives the minimum admissible value of n as 6.

Example...

The search is made as follows. The first two experiments are made at

$$x_1 = \frac{L_0}{2} - \frac{\delta}{2} = 0.5 - 0.0005 = 0.4995$$

$$x_2 = \frac{L_0}{2} + \frac{\delta}{2} = 0.5 + 0.0005 = 0.5005$$

with the function values given by

$$f_1 = f(x_1) = 0.4995(-1.0005) \simeq -0.49975$$

$$f_2 = f(x_2) = 0.5005(-0.9995) \simeq -0.50025$$

Since $f_2 < f_1$, the new interval of uncertainty will be (0.4995, 1.0). The second pa
of experiments is conducted at

$$x_3 = \left(0.4995 + \frac{1.0 - 0.4995}{2}\right) - 0.0005 = 0.74925$$

$$x_4 = \left(0.4995 + \frac{1.0 - 0.4995}{2}\right) + 0.0005 = 0.75025$$

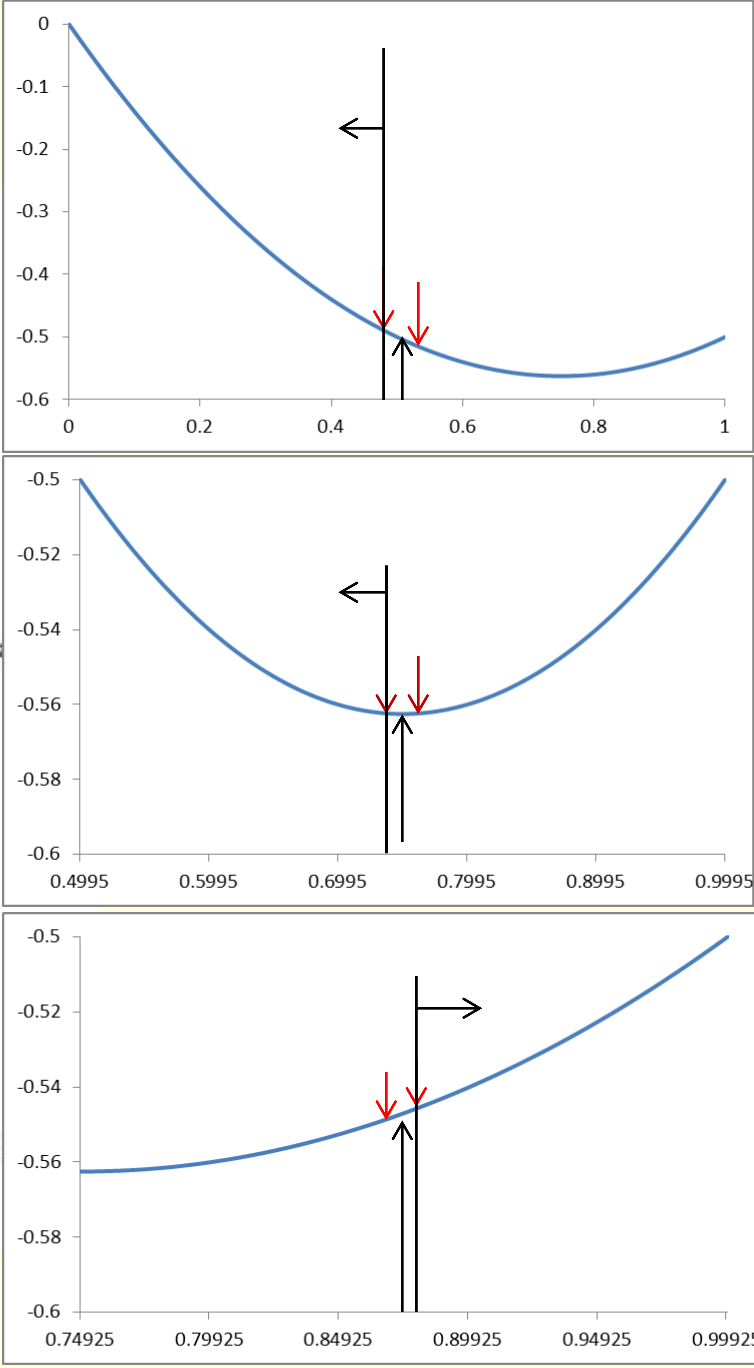
which give the function values as

$$f_3 = f(x_3) = 0.74925(-0.75075) = -0.5624994375$$

$$f_4 = f(x_4) = 0.75025(-0.74975) = -0.5624999375$$

Since $f_3 > f_4$, we delete (0.4995, x_3) and obtain the new interval of uncertainty as

$$(x_3, 1.0) = (0.74925, 1.0)$$



Example…

The final set of experiments will be conducted at

$$x_5 = \left(0.74925 + \frac{1.0 - 0.74925}{2}\right) - 0.0005 = 0.874125$$

$$x_6 = \left(0.74925 + \frac{1.0 - 0.74925}{2}\right) + 0.0005 = 0.875125$$

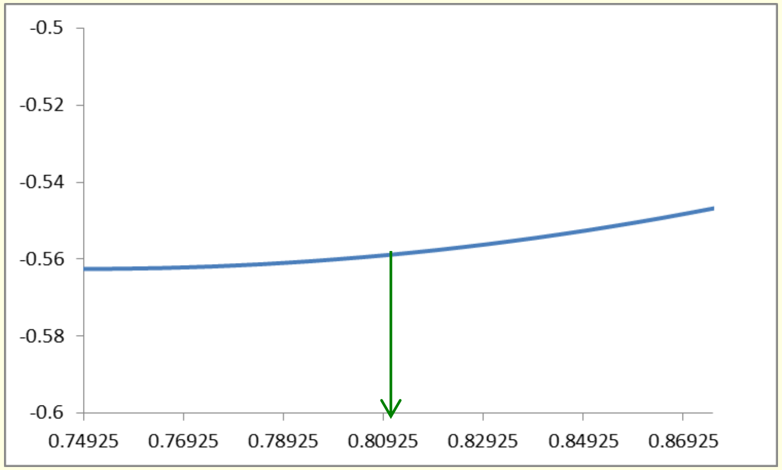
The corresponding function values are

$$f_5 = f(x_5) = 0.874125(-0.625875) = -0.5470929844$$

$$f_6 = f(x_6) = 0.875125(-0.624875) = -0.5468437342$$

Since $f_5 < f_6$, the new interval of uncertainty is given by $(x_3, x_6) = (0.74925, 0.875125)$. The middle point of this interval can be taken as optimum, and hence

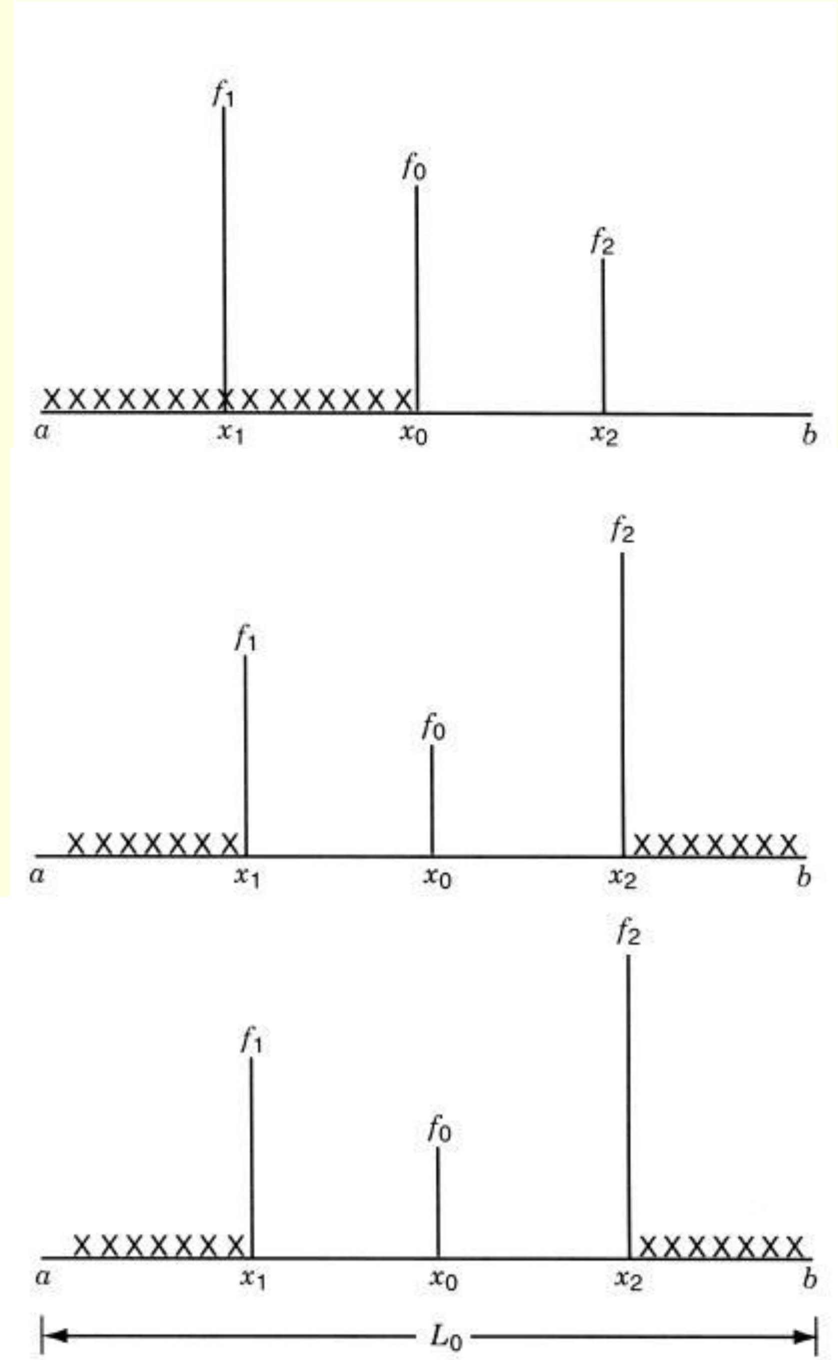
$$x_{opt} \simeq 0.8121875 \quad \text{and} \quad f_{opt} \simeq -0.5586327148$$



Interval Halving

- **Divide the initial interval into 4 equal parts using 3 points – one mid point (x_0) and 2 quarter points (x_1, x_2)**
- **Based on these function values, certain area (50%) [x_1, x_2] is eliminated**
- **2 points are created again using one of the existing points and process repeated again**
- **Interval of uncertainty remaining at the end of n experiments (n≥3 & n odd)**

$$L_n = \left(\frac{1}{2}\right)^{(n-1)/2} L_0$$



Interval Halving Example

Example 5.6 Find the minimum of $f = x(x - 1.5)$ in the interval (0.0, 1.0) to within 10% of the exact value.

SOLUTION If the middle point of the final interval of uncertainty is taken as the optimum point, the specified accuracy can be achieved if

$$\frac{1}{2}L_n \le \frac{L_0}{10} \quad \text{or} \quad \left(\frac{1}{2}\right)^{(n-1)/2} L_0 \le \frac{L_0}{5} \tag{E$_1$}$$

Since $L_0 = 1$, Eq. (E_1) gives

$$\frac{1}{2^{(n-1)/2}} \le \frac{1}{5} \quad \text{or} \quad 2^{(n-1)/2} \ge 5 \tag{E$_2$}$$

Since n has to be odd, inequality (E_2) gives the minimum permissible value of n as 7. With this value of $n = 7$, the search is conducted as follows. The first three experiments are placed at one-fourth points of the interval $L_0 = [a = 0, b = 1]$ as

$$x_1 = 0.25, \qquad f_1 = 0.25(-1.25) = -0.3125$$

$$x_0 = 0.50, \qquad f_0 = 0.50(-1.00) = -0.5000$$

$$x_2 = 0.75, \qquad f_2 = 0.75(-0.75) = -0.5625$$

Example...

Since $f_1 > f_0 > f_2$, we delete the interval $(a, x_0) = (0.0, 0.5)$, label x_2 and x_0 as the new x_0 and a so that $a = 0.5$, $x_0 = 0.75$, and $b = 1.0$. By dividing the new interval of uncertainty, $L_3 = (0.5, 1.0)$ into four equal parts, we obtain

$$x_1 = 0.625, \qquad f_1 = 0.625(-0.875) = -0.546875$$
$$x_0 = 0.750, \qquad f_0 = 0.750(-0.750) = -0.562500$$
$$x_2 = 0.875, \qquad f_2 = 0.875(-0.625) = -0.546875$$

Since $f_1 > f_0$ and $f_2 > f_0$, we delete both the intervals (a, x_1) and (x_2, b), and label x_1, x_0, and x_2 as the new a, x_0, and b, respectively. Thus the new interval of uncertainty will be $L_5 = (0.625, 0.875)$. Next, this interval is divided into four equal parts to obtain

$$x_1 = 0.6875, \qquad f_1 = 0.6875(-0.8125) = -0.558594$$
$$x_0 = 0.75, \qquad f_0 = 0.75(-0.75) = -0.5625$$
$$x_2 = 0.8125, \qquad f_2 = 0.8125(-0.6875) = -0.558594$$

Again we note that $f_1 > f_0$ and $f_2 > f_0$ and hence we delete both the intervals (a, x_1) and (x_2, b) to obtain the new interval of uncertainty as $L_7 = (0.6875, 0.8125)$. By taking the middle point of this interval (L_7) as optimum, we obtain

$$x_{\text{opt}} \approx 0.75 \quad \text{and} \quad f_{\text{opt}} \approx -0.5625$$

(This solution happens to be the exact solution in this case.)

Fibonacci Search

- **50% area eliminated in every iteration using 2 function evaluation $\Rightarrow$ 25% area eliminated per function evaluation**
- **Can we eliminate this much or more area using one function evaluation in each iteration?**
- **Technique uses Fibonacci numbers to carry out this area elimination process using one function evaluation in one iteration**

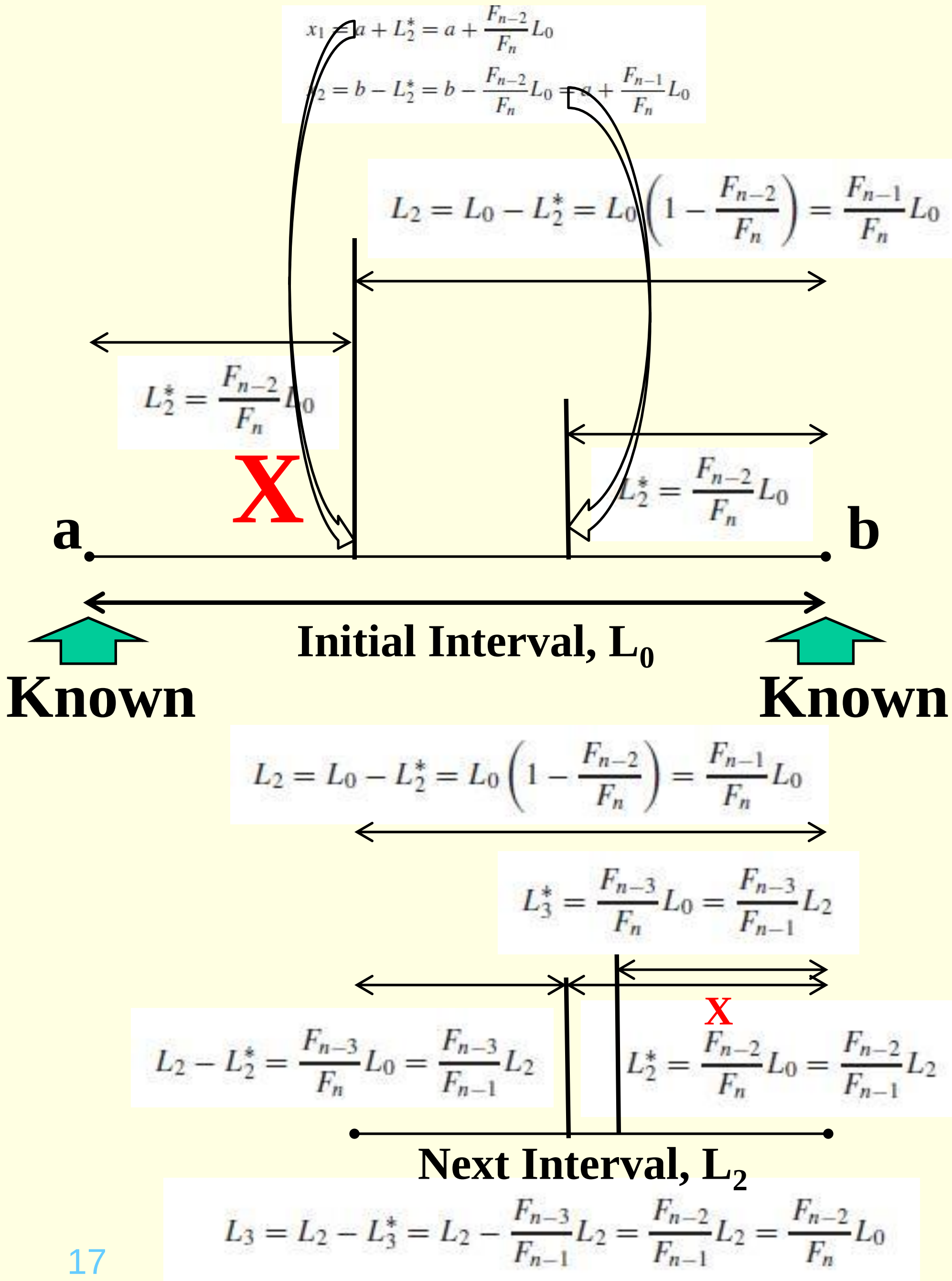
$$F_0 = F_1 = 1$$

$$F_n = F_{n-1} + F_{n-2}, \qquad n = 2, 3, 4, \ldots$$

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \ldots$$

Fibonacci Search

- Based on accuracy required, n and F_n are determined, n is total number of experiments
- 2 test points (x_1, x_2) are placed at a distance L_2^* from each end of L_0
- Using unimodality assumption, discard area
- In the remaining area, one experiment is already present, one needs to be introduced



$$L_j^* = \frac{F_{n-j}}{F_{n-(j-2)}} L_{j-1}$$

$$L_j = \frac{F_{n-(j-1)}}{F_n} L_0$$

$$\frac{L_j}{L_0} = \frac{F_{n-(j-1)}}{F_n}$$

$$\frac{L_n}{L_0} = \frac{F_1}{F_n} = \frac{1}{F_n}$$

Fibonacci Example

Example 5.7 Minimize $f(x) = 0.65 - [0.75/(1 + x^2)] - 0.65x \tan^{-1}(1/x)$ in the interval [0,3] by the Fibonacci method using $n = 6$. (Note that this objective is equivalent to the one stated in Example 5.2.)

SOLUTION Here $n = 6$ and $L_0 = 3.0$, which yield

$$L_2^* = \frac{F_{n-2}}{F_n} L_0 = \frac{5}{13}(3.0) = 1.153846$$

Thus the positions of the first two experiments are given by $x_1 = 1.153846$ and $x_2 = 3.0 - 1.153846 = 1.846154$ with $f_1 = f(x_1) = -0.207270$ and $f_2 = f(x_2) = -0.115843$. Since f_1 is less than f_2, we can delete the interval $[x_2, 3.0]$ by using the unimodality assumption (Fig. 5.10*a*). The third experiment is placed at $x_3 = 0 + (x_2 - x_1) = 1.846154 - 1.153846 = 0.692308$, with the corresponding function value of $f_3 = -0.291364$.

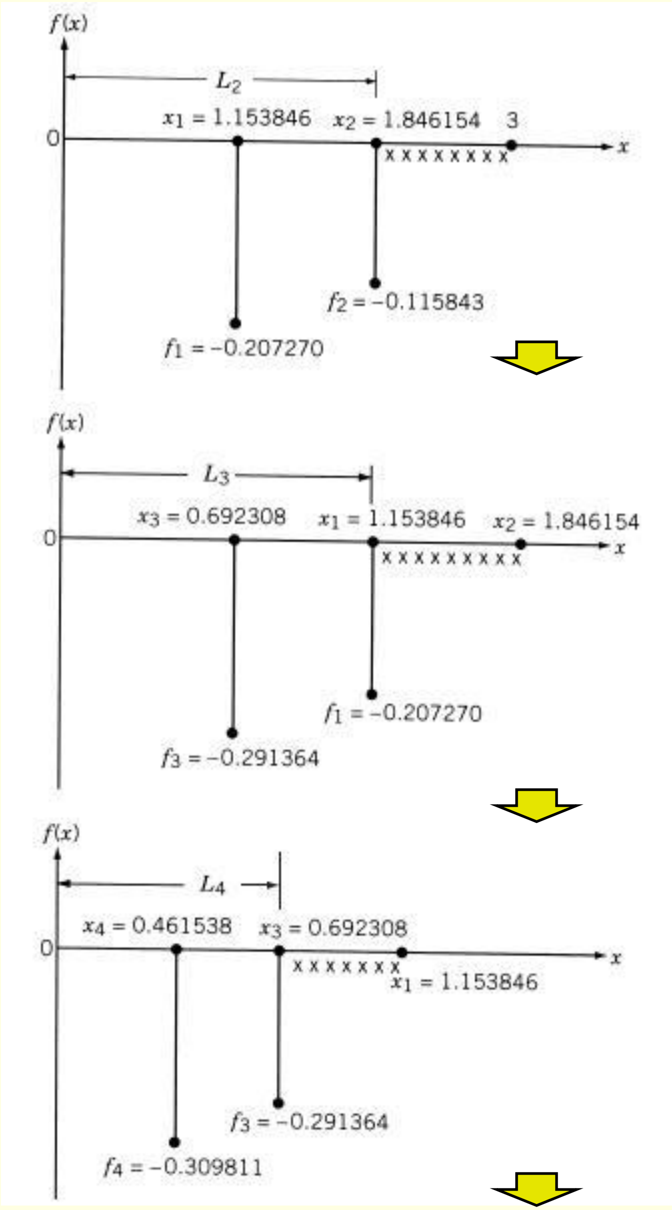
Since $f_1 > f_3$, we delete the interval $[x_1, x_2]$ (Fig. 5.10*b*). The next experiment is located at $x_4 = 0 + (x_1 - x_3) = 1.153846 - 0.692308 = 0.461538$ with $f_4 = -0.309811$. Nothing that $f_4 < f_3$, we delete the interval $[x_3, x_1]$ (Fig. 5.10*c*). The location of the next experiment can be obtained as $x_5 = 0 + (x_3 - x_4) = 0.692308 - 0.461538 = 0.230770$ with the corresponding objective function value of $f_5 = -0.263678$. Since $f_5 > f_4$, we delete the interval $[0, x_5]$ (Fig. 5.10*d*). The final experiment is positioned at $x_6 = x_5 + (x_3 - x_4) = 0.230770 + (0.692308 - 0.461538) = 0.461540$ with $f_6 = -0.309810$. (Note that, theoretically, the value of x_6 should be same as that of x_4; however, it is slightly different from x_4, due to round-off error).

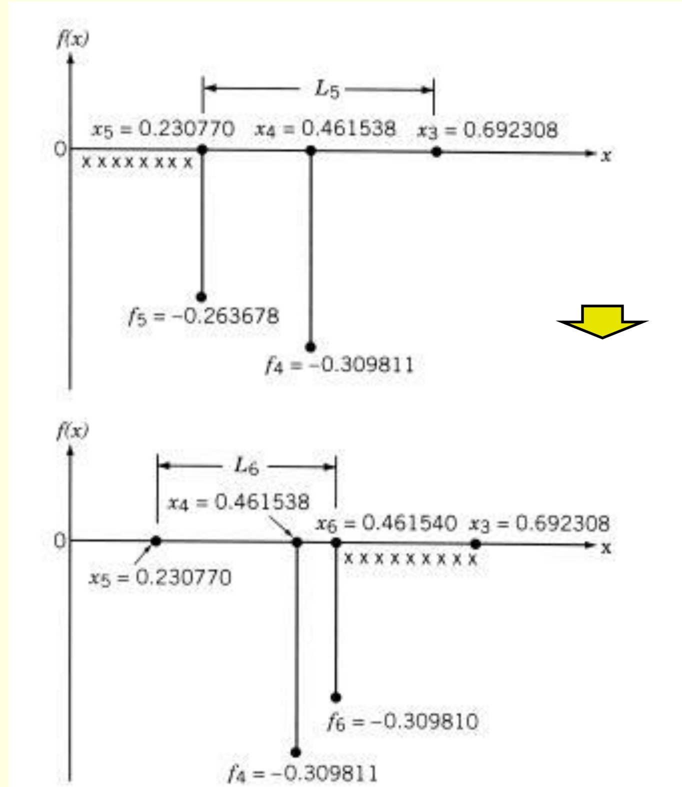
Since $f_6 > f_4$, we delete the interval $[x_6, x_3]$ and obtain the final interval of uncertainty as $L_6 = [x_5, x_6] = [0.230770, 0.461540]$ (Fig. 5.10*e*). The ratio of the final to the initial interval of uncertainty is

$$\frac{L_6}{L_0} = \frac{0.461540 - 0.230770}{3.0} = 0.076923$$

This value can be compared with Eq. (5.15), which states that if n experiments ($n = 6$) are planned, a resolution no finer than $1/F_n = 1/F_6 = \frac{1}{13} = 0.076923$ can be expected from the method.

Fibonacci Example





Golden Section Search

- Same as Fibonacci method except
  - Number of experiments need not to be mentioned in the beginning
  - Location of first 2 experiments does not need the information of total number of experiments – in this case we assume we are going to conduct a large number of experiments

$$L_2 = \lim_{N\to\infty} \frac{F_{N-1}}{F_N} L_0$$

$$L_3 = \lim_{N\to\infty} \frac{F_{N-2}}{F_N} L_0 = \lim_{N\to\infty} \frac{F_{N-2}}{F_{N-1}} \frac{F_{N-1}}{F_N} L_0$$

$$\simeq \lim_{N\to\infty} \left(\frac{F_{N-1}}{F_N}\right)^2 L_0$$

Value of N	2	3	4	5	6	7	8	9	10	∞
Ratio $\frac{F_{N-1}}{F_N}$	0.5	0.667	0.6	0.625	0.6156	0.619	0.6177	0.6181	0.6184	0.618

$$L_k = \lim_{N\to\infty} \left(\frac{F_{N-1}}{F_N}\right)^{k-1} L_0$$

$$F_N = F_{N-1} + F_{N-2}$$

$$\frac{F_N}{F_{N-1}} = 1 + \frac{F_{N-2}}{F_{N-1}}$$

$$\gamma = \lim_{N\to\infty} \frac{F_N}{F_{N-1}}$$

$$\gamma \simeq \frac{1}{\gamma} + 1$$

$$\gamma^2 - \gamma - 1 = 0$$

$\gamma = 1.618$,

$$L_k = \left(\frac{1}{\gamma}\right)^{k-1} L_0 = (0.618)^{k-1} L_0$$

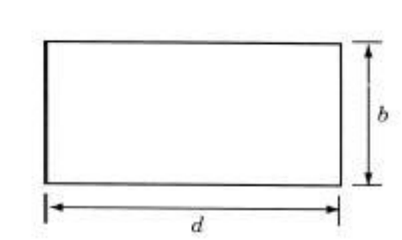
Golden Section Search

- Same as Fibonacci method except
  - First 2 experiments are positioned by

$$L_2^* = \frac{F_{N-2}}{F_N} L_0 = \frac{F_{N-2}}{F_{N-1}} \frac{F_{N-1}}{F_N} L_0 = \frac{L_0}{\gamma^2} = 0.382 L_0$$

  - Desired accuracy can stop the procedure

$$\frac{d+b}{d} = \frac{d}{b} = \gamma$$



Golden Section Example

Example 5.8 Minimize the function

$$f(x) = 0.65 - [0.75/(1 + x^2)] - 0.65x \tan^{-1}(1/x)$$

using the golden section method with $n = 6$.

SOLUTION The locations of the first two experiments are defined by $L_2^* = 0.382L_0 = (0.382)(3.0) = 1.1460$. Thus $x_1 = 1.1460$ and $x_2 = 3.0 - 1.1460 = 1.8540$ with $f_1 = f(x_1) = -0.208654$ and $f_2 = f(x_2) = -0.115124$. Since $f_1 < f_2$, we delete the interval $[x_2, 3.0]$ based on the assumption of unimodality and obtain the new interval of uncertainty as $L_2 = [0, x_2] = [0.0, 1.8540]$. The third experiment is placed at $x_3 = 0 + (x_2 - x_1) = 1.8540 - 1.1460 = 0.7080$. Since $f_3 = -0.288943$ is smaller than $f_1 = -0.208654$, we delete the interval $[x_1, x_2]$ and obtain the new interval of uncertainty as $[0.0, x_1] = [0.0, 1.1460]$. The position of the next experiment is given by $x_4 = 0 + (x_1 - x_3) = 1.1460 - 0.7080 = 0.4380$ with $f_4 = -0.308951$.

Since $f_4 < f_3$, we delete $[x_3, x_1]$ and obtain the new interval of uncertainty as $[0, x_3] = [0.0, 0.7080]$. The next experiment is placed at $x_5 = 0 + (x_3 - x_4) = 0.7080 - 0.4380 = 0.2700$. Since $f_5 = -0.278434$ is larger than $f_4 = -0.308951$, we delete the interval $[0, x_5]$ and obtain the new interval of uncertainty as $[x_5, x_3] = [0.2700, 0.7080]$. The final experiment is placed at $x_6 = x_5 + (x_3 - x_4) = 0.2700 + (0.7080 - 0.4380) = 0.5400$ with $f_6 = -0.308234$. Since $f_6 > f_4$, we delete the interval $[x_6, x_3]$ and obtain the final interval of uncertainty as $[x_5, x_6] = [0.2700, 0.5400]$. Note that this final interval of uncertainty is slightly larger than the one found in the Fibonacci method, [0.461540, 0.230770]. The ratio of the final to the initial interval of uncertainty in the present case is

$$\frac{L_6}{L_0} = \frac{0.5400 - 0.2700}{3.0} = \frac{0.27}{3.0} = 0.09$$

Comparison

Method	Formula	$n = 5$	$n = 10$
Exhaustive search	$L_n = \frac{2}{n+1} L_0$	$0.33333L_0$	$0.18182L_0$
Dichotomous search ($\delta = 0.01$ and n = even)	$L_n = \frac{L_0}{2^{n/2}} + \delta\left(1 - \frac{1}{2^{n/2}}\right)$	$\frac{1}{4}L_0 + 0.0075$ with $n = 4$, $\frac{1}{8}L_0 + 0.00875$ with $n = 6$	$0.03125L_0 + 0.0096875$
Interval halving ($n \geq 3$ and odd)	$L_n = (\frac{1}{2})^{(n-1)/2} L_0$	$0.25L_0$	$0.0625L_0$ with $n = 9$, $0.03125L_0$ with $n = 11$
Fibonacci	$L_n = \frac{1}{F_n} L_0$	$0.125L_0$	$0.01124L_0$
Golden section	$L_n = (0.618)^{n-1} L_0$	$0.1459L_0$	$0.01315L_0$

Method	Error: $\frac{1}{2}\frac{L_n}{L_0} \leq 0.1$	Error: $\frac{1}{2}\frac{L_n}{L_0} \leq 0.01$
Exhaustive search	$n \geq 9$	$n \geq 99$
Dichotomous search ($\delta = 0.01$, $L_0 = 1$)	$n \geq 6$	$n \geq 14$
Interval halving ($n \geq 3$ and odd)	$n \geq 7$	$n \geq 13$
Fibonacci	$n \geq 4$	$n \geq 9$
Golden section	$n \geq 5$	$n \geq 10$

Quadratic Interpolation

- **Uses Function values, no derivatives**
- **Useful for cases when derivative computation is not favorable**

3 Stage approach

- **Stage 1: Normalize the direction vector**
- **Stage 2: Apply a quadratic approximation to the given function and find the minimum of the given function through successive quadratic approximation approach**
- **Stage 3: Terminate based on different criteria**

QI (contd.)

Stage 1: Direction vector normalization

- **Any n dimensional direction vector s = {s_1, s_2,… s_i, …, s_n} can be normalized by dividing each component by Δ**

$$\Delta = \underset{i}{\mathrm{Max}} \left| s_i \right|$$

- **Other way:**

$$\Delta = \sqrt{\left(s_1^2 + s_2^2 + \ldots + s_n^2 \right)}$$

QI (contd.)

Stage 2: Quadratic approximation

- **$f(\lambda)$ is the univariate function to which the $h(\lambda)$ quadratic function approximation needs to be fitted**

$$h(\lambda) = a + b\lambda + c\lambda^2$$

- **We need 3 points (A, B and C) to find coefficients for this function**

$$f_A = a + bA + cA^2$$
$$f_B = a + bB + cB^2$$
$$f_C = a + bC + cC^2$$

$$a = \frac{f_A BC(C-B) + f_B CA(A-C) + f_C AB(B-A)}{(A-B)(B-C)(C-A)}$$

$$b = \frac{f_A(B^2-C^2) + f_B(C^2-A^2) + f_C(A^2-B^2)}{(A-B)(B-C)(C-A)}$$

$$c = -\frac{f_A(B-C) + f_B(C-A) + f_C(A-B)}{(A-B)(B-C)(C-A)}$$

QI (contd.)

Quadratic approximation (by making $h'(\lambda) = 0$)

$$\tilde{\lambda}^* = \frac{-b}{2c} = \frac{f_A(B^2 - C^2) + f_B(C^2 - A^2) + f_C(A^2 - B^2)}{2[f_A(B - C) + f_B(C - A) + f_C(A - B)]}$$

- **Assuming 3 points (A, B and C) as $\lambda = 0$ (f_A), t (f_B) & 2t (f_C), where t is a trial step to be assumed ($\lambda = 0$ saves one function evaluation – next iteration onwards)**

$$a = f_A$$

$$b = \frac{4f_B - 3f_A - f_C}{2t}$$

$$c = \frac{f_C + f_A - 2f_B}{2t^2}$$

$$\tilde{\lambda}^* = \frac{4f_B - 3f_A - f_C}{4f_B - 2f_C - 2f_A}t$$

provided $h''(\lambda) > 0$

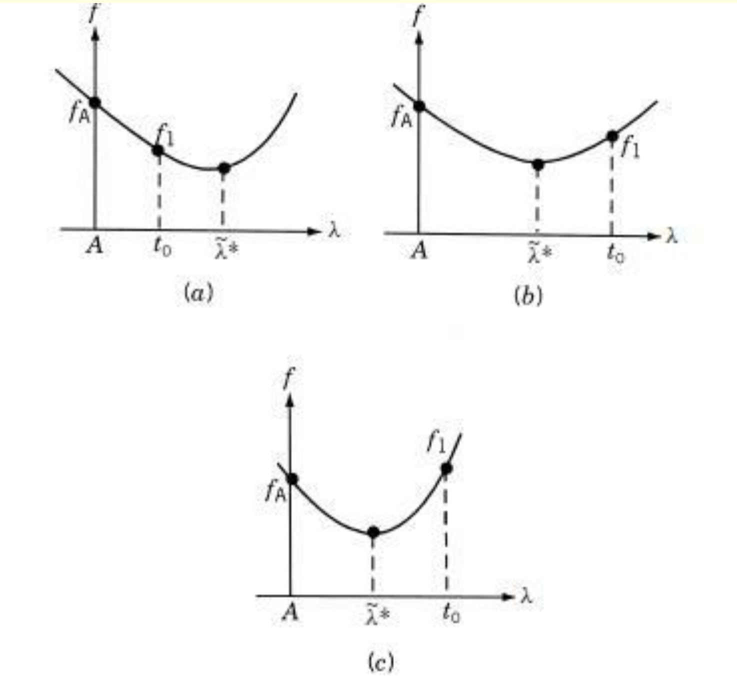
$$c = \frac{f_C + f_A - 2f_B}{2t^2} > 0$$

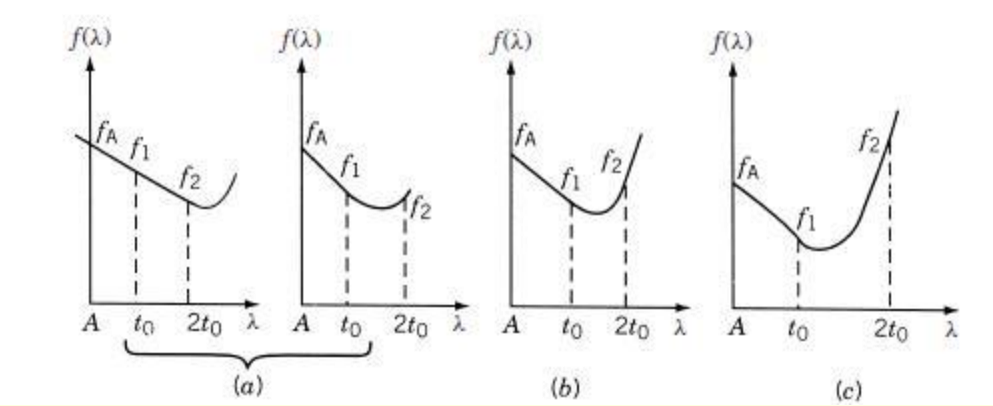
$$\frac{f_A + f_C}{2} > f_B$$

QI (contd.)

This can be ensured by

- Choose $f_A = f(\lambda = 0)$ and compute $f_1 = f(\lambda = t_0)$
- If $f_1 > f_A$, $f_C = f_1$, compute $f_B = f(\lambda = t_0/2)$, compute optimum λ using $t = t_0/2$
- If $f_1 < f_A$, $f_B = f_1$, compute $f_2 = f(\lambda = 2t_0)$
- If $f_2 > f_1$, $f_C = f_2$ and $f_B = f_1$, compute optimum λ using $t = t_0$
- If $f_2 < f_1$, $f_2 = f_1$ and $t = 2t_0$ and repeat above 3 steps till we find $f_2 > f_1$





QI (contd.)

Stage 3: Termination

- **We need to ensure that optimum λ value of approximate function h(λ) is sufficiently close to the true optimum λ value of original function f(λ)**

- **Termination criteria**

$$\left|\frac{h(\tilde{\lambda}^*) - f(\tilde{\lambda}^*)}{f(\tilde{\lambda}^*)}\right| \le \varepsilon_1$$

$$\left|\frac{f(\tilde{\lambda}^* + \Delta\tilde{\lambda}^*) - f(\tilde{\lambda}^* - \Delta\tilde{\lambda}^*)}{2\Delta\tilde{\lambda}^*}\right| \le \varepsilon_2$$

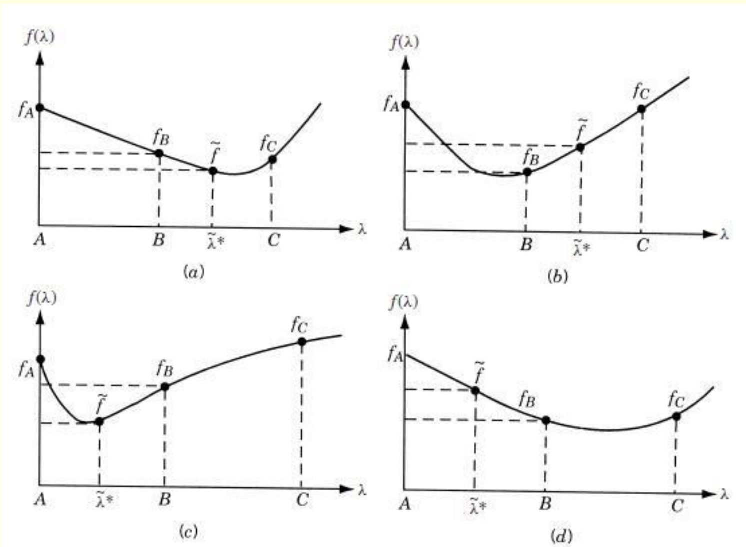
- **If termination criteria satisfied, stop**
- **Else refit a new quadratic polynomial using 3 best points out of 4 points from the previous iteration**

$$h'(\lambda) = a' + b'\lambda + c'\lambda^2$$

QI (contd.)

Refitting

- **Selection of 3 best points out of all possible situations**



		New points for refitting	
Case	Characteristics	New	Old
1	$\tilde{\lambda}^* > B$	A	B
	$\tilde{f} < f_B$	B	$\tilde{\lambda}^*$
		C	C
		Neglect old A	
2	$\tilde{\lambda}^* > B$	A	A
	$\tilde{f} > f_B$	B	B
		C	$\tilde{\lambda}^*$
		Neglect old C	
3	$\tilde{\lambda}^* < B$	A	A
	$\tilde{f} < f_B$	B	$\tilde{\lambda}^*$
		C	B
		Neglect old C	
4	$\tilde{\lambda}^* < B$	A	$\tilde{\lambda}^*$
	$\tilde{f} > f_B$	B	B
		C	C
		Neglect old A	

QI - Example

Example 5.10 Find the minimum of $f = \lambda^5 - 5\lambda^3 - 20\lambda + 5$.

SOLUTION Since this is not a multivariable optimization problem, we can proceed directly to stage 2. Let the initial step size be taken as $t_0 = 0.5$ and $A = 0$.

Iteration 1

$$f_A = f(\lambda = 0) = 5$$

$$f_1 = f(\lambda = t_0) = 0.03125 - 5(0.125) - 20(0.5) + 5 = -5.59375$$

Since $f_1 < f_A$, we set $f_B = f_1 = -5.59375$, and find that

$$f_2 = f(\lambda = 2t_0 = 1.0) = -19.0$$

As $f_2 < f_1$, we set new $t_0 = 1$ and $f_1 = -19.0$. Again we find that $f_1 < f_A$ and hence set $f_B = f_1 = -19.0$, and find that $f_2 = f(\lambda = 2t_0 = 2) = -43$. Since $f_2 < f_1$, we again set $t_0 = 2$ and $f_1 = -43$. As this $f_1 < f_A$, set $f_B = f_1 = -43$ and evaluate $f_2 = f(\lambda = 2t_0 = 4) = 629$. This time $f_2 > f_1$ and hence we set $f_C = f_2 = 629$ and compute $\tilde{\lambda}^*$ from Eq. (5.40) as

$$\tilde{\lambda}^* = \frac{4(-43) - 3(5) - 629}{4(-43) - 2(629) - 2(5)}(2) = \frac{1632}{1440} = 1.135$$

Convergence test: Since $A = 0$, $f_A = 5$, $B = 2$, $f_B = -43$, $C = 4$, and $f_C = 629$, the values of a, b, and c can be found to be

$$a = 5, \quad b = -204, \quad c = 90$$

QI - Example

and

$$h(\tilde{\lambda}^*) = h(1.135) = 5 - 204(1.135) + 90(1.135)^2 = -110.9$$

Since

$$\tilde{f} = f(\tilde{\lambda}^*) = (1.135)^5 - 5(1.135)^3 - 20(1.135) + 5.0 = -23.127$$

we have

$$\left|\frac{h(\tilde{\lambda}^*) - f(\tilde{\lambda}^*)}{f(\tilde{\lambda}^*)}\right| = \left|\frac{-116.5 + 23.127}{-23.127}\right| = 3.8$$

As this quantity is very large, convergence is not achieved and hence we have to use *refitting*.

QI - Example

Iteration 2

Since $\tilde{\lambda}^* < B$ and $\tilde{f} > f_B$, we take the new values of A, B, and C as

$$A = 1.135, \qquad f_A = -23.127$$
$$B = 2.0, \qquad f_B = -43.0$$
$$C = 4.0, \qquad f_C = 629.0$$

and compute new $\tilde{\lambda}^*$, using Eq. (5.36), as

$$\tilde{\lambda}^* = \frac{(-23.127)(4.0 - 16.0) + (-43.0)(16.0 - 1.29) + (629.0)(1.29 - 4.0)}{2[(-23.127)(2.0 - 4.0) + (-43.0)(4.0 - 1.135) + (629.0)(1.135 - 2.0)]} = 1.661$$

Convergence test: To test the convergence, we compute the coefficients of the quadratic as

$$a = 288.0, \quad b = -417.0, \quad c = 125.3$$

As

$$h(\tilde{\lambda}^*) = h(1.661) = 288.0 - 417.0(1.661) + 125.3(1.661)^2 = -59.7$$
$$\tilde{f} = f(\tilde{\lambda}^*) = 12.8 - 5(4.59) - 20(1.661) + 5.0 = -38.37$$

we obtain

$$\left| \frac{h(\tilde{\lambda}^*) - f(\tilde{\lambda}^*)}{f(\tilde{\lambda}^*)} \right| = \left| \frac{-59.70 + 38.37}{-38.37} \right| = 0.556$$

Since this quantity is not sufficiently small, we need to proceed to the next refit.

Cubic Interpolation

- **Uses Function values and derivatives**

4 Stage approach

- **Stage 1: Normalize the direction vector**
- **Stage 2: Bracket the optimum point**
- **Stage 3: Apply a cubic approximation to the given function and find the minimum of the given function through successive cubic approximation approach**
- **Stage 4: Terminate based on different criteria**

CI (contd.)

Stage 1: Direction vector normalization

- **Any n dimensional direction vector s = $\{s_1, s_2, \ldots s_i, \ldots, s_n\}$ can be normalized by dividing each component by Δ**

$$\Delta = \operatorname*{Max}_{i} |s_i|$$

- **Other way:**

$$\Delta = \sqrt{\left(s_1^2 + s_2^2 + \ldots + s_n^2\right)}$$

CI (contd.)

Stage 2: **Bracketing optimum**

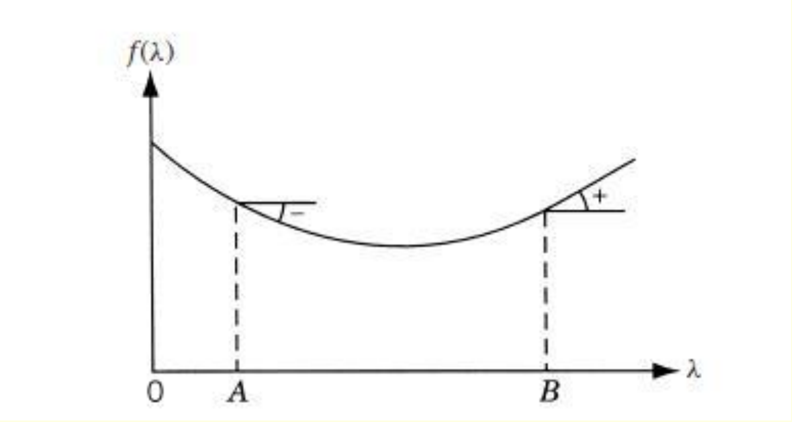
- **$f(\lambda)$ is univariate function – to bracket the optimum, derivative information at 2 points checked for signs (one –ve and one +ve)**

$$f'(\lambda) = \frac{df}{d\lambda} = \frac{d}{d\lambda} f(\mathbf{X} + \lambda \mathbf{S}) = \mathbf{S}^{\mathrm{T}} \nabla f(\mathbf{X} + \lambda \mathbf{S})$$

- **At $\lambda = 0$ (point A), since S is assumed to be the direction of descent**

$$\left.\frac{df}{d\lambda}\right|_{\lambda=0} = \mathbf{S}^{\mathrm{T}} \nabla f(\mathbf{X}) < 0$$

- **We find one more point (point B) where the slope ($df/d\lambda$) is +ve - $\lambda + t_0$, $2t_0$, $4t_0$, $8t_0$ etc. till the above condition satisfied**



CI (contd.)

Stage 3: **Cubic approximation**

- **f(λ) is the univariate function to which the h(λ) cubic function approximation needs to be fitted**

$$h(\lambda) = a + b\lambda + c\lambda^2 + \lambda^3$$

- **We need 4 data (function value of A, B & derivative information at A, B) to find coefficients**

$$f_A = a + bA + cA^2 + dA^3$$
$$f_B = a + bB + cB^2 + dB^3$$
$$f'_A = b + 2cA + 3dA^2$$
$$f'_B = b + 2cB + 3dB^2$$

$$a = f_A - bA - cA^2 - dA^3$$

$$b = \frac{1}{(A-B)^2}(B^2 f'_A + A^2 f'_B + 2ABZ)$$

$$c = -\frac{1}{(A-B)^2}[(A+B)Z + Bf'_A + Af'_B]$$

$$d = \frac{1}{3(A-B)^2}(2Z + f'_A + f'_B)$$

$$Z = \frac{3(f_A - f_B)}{B-A} + f'_A + f'_B$$

CI (contd.)

Stage 3: **Cubic approximation**

- **Application of optimization conditions leads to**

$$\tilde{\lambda}^* = A + \frac{f'_A + Z \pm Q}{f'_A + f'_B + 2Z}(B - A)$$

$$Q = (Z^2 - f'_A f'_B)^{1/2}$$

A = 0

$$\tilde{\lambda}^* = B\frac{f'_A + Z \pm Q}{f'_A + f'_B + 2Z}$$

$$Q = (Z^2 - f'_A f'_B)^{1/2} > 0$$

$$Z = \frac{3(f_A - f_B)}{B} + f'_A + f'_B$$

CI (contd.)

Stage 4: **Termination**

- We need to ensure that optimum λ value of approximate function h(λ) is sufficiently close to the true optimum λ value of original function f(λ)
- Termination criteria

$$\left| \frac{h(\tilde{\lambda}^*) - f(\tilde{\lambda}^*)}{f(\tilde{\lambda}^*)} \right| \le \varepsilon_1$$

$$\left| \frac{df}{d\lambda} \right|_{\tilde{\lambda}^*} = |\mathbf{S}^{\mathrm{T}} \nabla f|_{\tilde{\lambda}^*}| \le \varepsilon_2$$

$$\left| \frac{\mathbf{S}^{\mathrm{T}} \nabla f}{|\mathbf{S}||\nabla f|} \right|_{\tilde{\lambda}^*} \le \varepsilon_2$$

- If termination criteria satisfied, stop
- Else refit a new cubic polynomial using 2 best points out of 3 points from the previous iteration

$$h'(\lambda) = a' + b'\lambda + c'\lambda^2 + d'\lambda^3$$

CI - Example

Example 5.11 Find the minimum of $f = \lambda^5 - 5\lambda^3 - 20\lambda + 5$ by the cubic interpolation method.

SOLUTION Since this problem has not arisen during a multivariable optimization process, we can skip stage 1. We take $A = 0$ and find that

$$\frac{df}{d\lambda}(\lambda = A = 0) = 5\lambda^4 - 15\lambda^2 - 20\Big|_{\lambda=0} = -20 < 0$$

To find B at which $df/d\lambda$ is nonnegative, we start with $t_0 = 0.4$ and evaluate the derivative at $t_0, 2t_0, 4t_0, \ldots$. This gives

$$f'(t_0 = 0.4) = 5(0.4)^4 - 15(0.4)^2 - 20.0 = -22.272$$

$$f'(2t_0 = 0.8) = 5(0.8)^4 - 15(0.8)^2 - 20.0 = -27.552$$

$$f'(4t_0 = 1.6) = 5(1.6)^4 - 15(1.6)^2 - 20.0 = -25.632$$

$$f'(8t_0 = 3.2) = 5(3.2)^4 - 15(3.2)^2 - 20.0 = 350.688$$

Thus we find that†

$$A = 0.0, \qquad f_A = 5.0, \qquad f'_A = -20.0$$

$$B = 3.2, \qquad f_B = 113.0, \qquad f'_B = 350.688$$

$$A < \lambda^* < B$$

CI - Example

Iteration 1

To find the value of $\tilde{\lambda}^*$ and to test the convergence criteria, we first compute Z and Q as

$$Z = \frac{3(5.0 - 113.0)}{3.2} - 20.0 + 350.688 = 229.588$$

$$Q = [229.588^2 + (20.0)(350.688)]^{1/2} = 244.0$$

Hence

$$\tilde{\lambda}^* = 3.2\left(\frac{-20.0 + 229.588 \pm 244.0}{-20.0 + 350.688 + 459.176}\right) = 1.84 \quad \text{or} \quad -0.1396$$

By discarding the negative value, we have

$$\tilde{\lambda}^* = 1.84$$

Convergence criterion: If $\tilde{\lambda}^*$ is close to the true minimum, λ^*, then $f'(\tilde{\lambda}^*) = df(\tilde{\lambda}^*)/d\lambda$ should be approximately zero. Since $f' = 5\lambda^4 - 15\lambda^2 - 20$,

$$f'(\tilde{\lambda}^*) = 5(1.84)^4 - 15(1.84)^2 - 20 = -13.0$$

Since this is not small, we go to the next iteration or refitting. As $f'(\tilde{\lambda}^*) < 0$, we take $A = \tilde{\lambda}^*$ and

$$f_A = f(\tilde{\lambda}^*) = (1.84)^5 - 5(1.84)^3 - 20(1.84) + 5 = -41.70$$

Thus

$$A = 1.84, \qquad f_A = -41.70, \qquad f'_A = -13.0$$

$$B = 3.2, \qquad f_B = 113.0, \qquad f'_B = 350.688$$

$$A < \tilde{\lambda}^* < B$$

CI - Example

Iteration 2

$$Z = \frac{3(-41.7 - 113.0)}{3.20 - 1.84} - 13.0 + 350.688 = -3.312$$

$$Q = [(-3.312)^2 + (13.0)(350.688)]^{1/2} = 67.5$$

Hence

$$\tilde{\lambda}^* = 1.84 + \frac{-13.0 - 3.312 \pm 67.5}{-13.0 + 350.688 - 6.624}(3.2 - 1.84) = 2.05$$

Convergence criterion:

$$f'(\tilde{\lambda}^*) = 5.0(2.05)^4 - 15.0(2.05)^2 - 20.0 = 5.35$$

Since this value is large, we go the next iteration with $B = \tilde{\lambda}^* = 2.05$ [as $f'(\tilde{\lambda}^*) > 0$] and

$$f_B = (2.05)^5 - 5.0(2.05)^3 - 20.0(2.05) + 5.0 = -42.90$$

Thus

$$A = 1.84, \quad f_A = -41.70, \quad f'_A = -13.00$$

$$B = 2.05, \quad f_B = -42.90, \quad f'_B = 5.35$$

$$A < \lambda^* < B$$

CI - Example

Iteration 3

$$Z = \frac{3.0(-41.70 + 42.90)}{(2.05 - 1.84)} - 13.00 + 5.35 = 9.49$$

$$Q = [(9.49)^2 + (13.0)(5.35)]^{1/2} = 12.61$$

Therefore,

$$\tilde{\lambda}^* = 1.84 + \frac{-13.00 + 9.49 \pm 12.61}{-13.00 + 5.35 + 18.98}(2.05 - 1.84) = 2.0086$$

Convergence criterion:

$$f'(\tilde{\lambda}^*) = 5.0(2.0086)^4 - 15.0(2.0086)^2 - 20.0 = 0.855$$

Assuming that this value is close to zero, we can stop the iterative process and take

$$\lambda^* \simeq \tilde{\lambda}^* = 2.0086$$

Root Finding Methods

- **Necessary condition to find the minimum of function $f(\lambda)$ is $f'(\lambda) = 0$**
- **Solving $f'(\lambda) = 0$ is numerical analysis is done by ROOT FINDING METHODS e.g. Newton-Raphson, Quasi Newton, secant methods etc. which is synonymous to finding the minimum**

Newton Raphson

- **Perform a quadratic approximation for any given function f(λ) around a point λ_i for which the minimum needs to be found**

$$f(\lambda) = f(\lambda_i) + f'(\lambda_i)(\lambda - \lambda_i) + \frac{1}{2} f''(\lambda_i)(\lambda - \lambda_i)^2$$

- **Applying the optimization criteria of first derivative going to 0**

$$f'(\lambda) = f'(\lambda_i) + f''(\lambda_i)(\lambda - \lambda_i) = 0$$

- **New points are generated using**

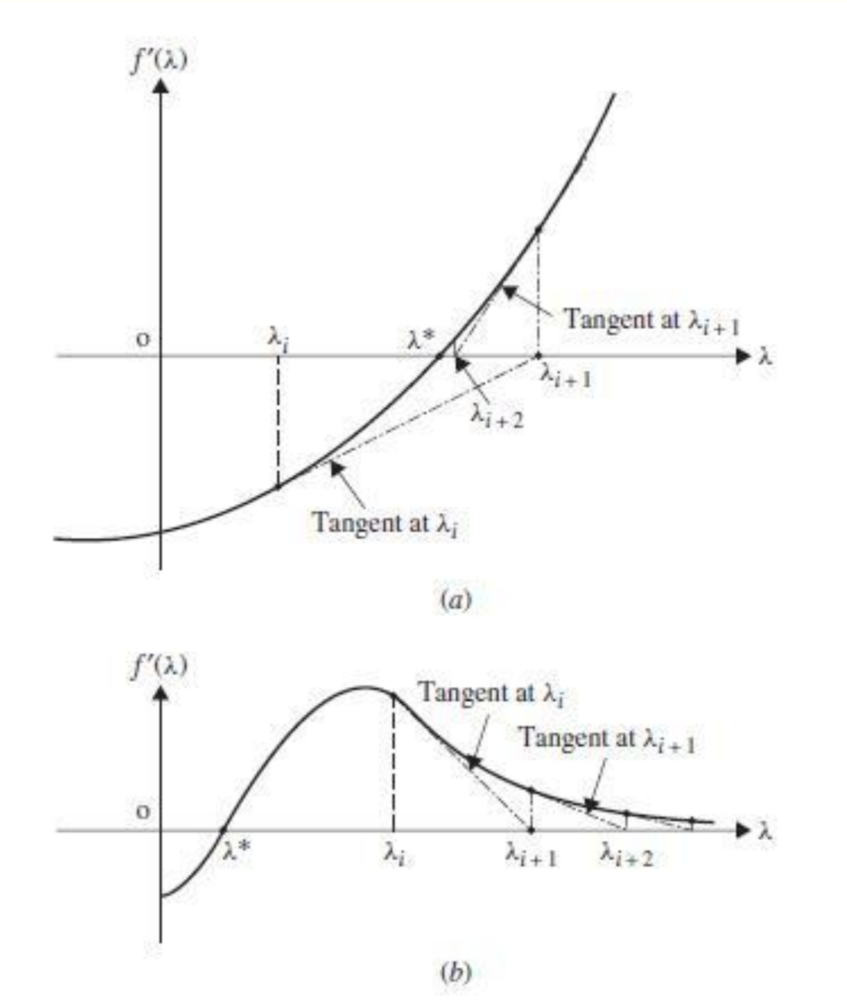
$$\lambda_{i+1} = \lambda_i - \frac{f'(\lambda_i)}{f''(\lambda_i)}$$

- **Terminate when**

$$|f'(\lambda_{i+1})| \leq \varepsilon$$

NR (contd.)

- Originally developed by Newton for solving nonlinear equations, modified by Raphson later
- Uses both, first and second order derivatives of the function f(λ)
- If $f''(\lambda) \neq 0$, NR has fastest convergence property, quadratic convergence
- If the initial solution is not sufficiently close to the true optimum, NR can diverge instead of converging



NR - Example

Example 5.12 Find the minimum of the function

$$f(\lambda) = 0.65 - \frac{0.75}{1+\lambda^2} - 0.65\lambda \tan^{-1}\frac{1}{\lambda}$$

using the Newton–Raphson method with the starting point $\lambda_1 = 0.1$. Use $\varepsilon = 0.01$ in Eq. (5.66) for checking the convergence.

SOLUTION The first and second derivatives of the function $f(\lambda)$ are given by

$$f'(\lambda) = \frac{1.5\lambda}{(1+\lambda^2)^2} + \frac{0.65\lambda}{1+\lambda^2} - 0.65\tan^{-1}\frac{1}{\lambda}$$

$$f''(\lambda) = \frac{1.5(1-3\lambda^2)}{(1+\lambda^2)^3} + \frac{0.65(1-\lambda^2)}{(1+\lambda^2)^2} + \frac{0.65}{1+\lambda^2} = \frac{2.8-3.2\lambda^2}{(1+\lambda^2)^3}$$

Iteration 1

$$\lambda_1 = 0.1, \quad f(\lambda_1) = -0.188197, \quad f'(\lambda_1) = -0.744832, \quad f''(\lambda_1) = 2.68659$$

$$\lambda_2 = \lambda_1 - \frac{f'(\lambda_1)}{f''(\lambda_1)} = 0.377241$$

Convergence check: $|f'(\lambda_2)| = |-0.138230| > \varepsilon$.

NR - Example

Iteration 2

$$f(\lambda_2) = -0.303279, \quad f'(\lambda_2) = -0.138230, \quad f''(\lambda_2) = 1.57296$$

$$\lambda_3 = \lambda_2 - \frac{f'(\lambda_2)}{f''(\lambda_2)} = 0.465119$$

Convergence check: $|f'(\lambda_3)| = |-0.0179078| > \varepsilon$.

Iteration 3

$$f(\lambda_3) = -0.309881, \quad f'(\lambda_3) = -0.0179078, \quad f''(\lambda_3) = 1.17126$$

$$\lambda_4 = \lambda_3 - \frac{f'(\lambda_3)}{f''(\lambda_3)} = 0.480409$$

Convergence check: $|f'(\lambda_4)| = |-0.0005033| < \varepsilon$.

Since the process has converged, the optimum solution is taken as $\lambda^* \approx \lambda_4 = 0.480409$.

Quasi Newton

- In case the function f(λ) is not available in closed form or computation of derivative is not possible analytically, we use the finite difference form of computing derivative (e.g. central difference – others could have been used)

$$f'(\lambda_i) = \frac{f(\lambda_i + \Delta\lambda) - f(\lambda_i - \Delta\lambda)}{2\Delta\lambda}$$

$$f''(\lambda_i) = \frac{f(\lambda_i + \Delta\lambda) - 2f(\lambda_i) + f(\lambda_i - \Delta\lambda)}{\Delta\lambda^2}$$

- Algorithm and convergence criteria becomes

$$\lambda_{i+1} = \lambda_i - \frac{\Delta\lambda[f(\lambda_i + \Delta\lambda) - f(\lambda_i - \Delta\lambda)]}{2[f(\lambda_i + \Delta\lambda) - 2f(\lambda_i) + f(\lambda_i - \Delta\lambda)]}$$

$$|f'(\lambda_{i+1})| = \left|\frac{f(\lambda_{i+1} + \Delta\lambda) - f(\lambda_{i+1} - \Delta\lambda)}{2\Delta\lambda}\right| \le \varepsilon$$

- Function evaluation required at $f(\lambda_i + \Delta)$, $f(\lambda_i - \Delta)$ apart from $f(\lambda_i)$

QN - Example

Example 5.13 Find the minimum of the function

$$f(\lambda) = 0.65 - \frac{0.75}{1+\lambda^2} - 0.65\lambda \tan^{-1}\frac{1}{\lambda}$$

using quasi-Newton method with the starting point $\lambda_1 = 0.1$ and the step size $\Delta\lambda = 0.01$ in central difference formulas. Use $\varepsilon = 0.01$ in Eq. (5.70) for checking the convergence.

SOLUTION

Iteration 1

$$\lambda_1 = 0.1, \quad \Delta\lambda = 0.01, \quad \varepsilon = 0.01, \quad f_1 = f(\lambda_1) = -0.188197,$$

$$f_1^+ = f(\lambda_1 + \Delta\lambda) = -0.195512, \quad f_1^- = f(\lambda_1 - \Delta\lambda) = -0.180615$$

$$\lambda_2 = \lambda_1 - \frac{\Delta\lambda(f_1^+ - f_1^-)}{2(f_1^+ - 2f_1 + f_1^-)} = 0.377882$$

Convergence check:

$$|f'(\lambda_2)| = \left|\frac{f_2^+ - f_2^-}{2\Delta\lambda}\right| = 0.137300 > \varepsilon$$

QN - Example

Iteration 2

$$f_2 = f(\lambda_2) = -0.303368, \quad f_2^+ = f(\lambda_2 + \Delta\lambda) = -0.304662,$$

$$f_2^- = f(\lambda_2 - \Delta\lambda) = -0.301916$$

$$\lambda_3 = \lambda_2 - \frac{\Delta\lambda(f_2^+ - f_2^-)}{2(f_2^+ - 2f_2 + f_2^-)} = 0.465390$$

Convergence check:

$$|f'(\lambda_3)| = \left|\frac{f_3^+ - f_3^-}{2\Delta\lambda}\right| = 0.017700 > \varepsilon$$

Iteration 3

$$f_3 = f(\lambda_3) = -0.309885, \quad f_3^+ = f(\lambda_3 + \Delta\lambda) = -0.310004,$$

$$f_3^- = f(\lambda_3 - \Delta\lambda) = -0.309650$$

$$\lambda_4 = \lambda_3 - \frac{\Delta\lambda(f_3^+ - f_3^-)}{2(f_3^+ - 2f_3 + f_3^-)} = 0.480600$$

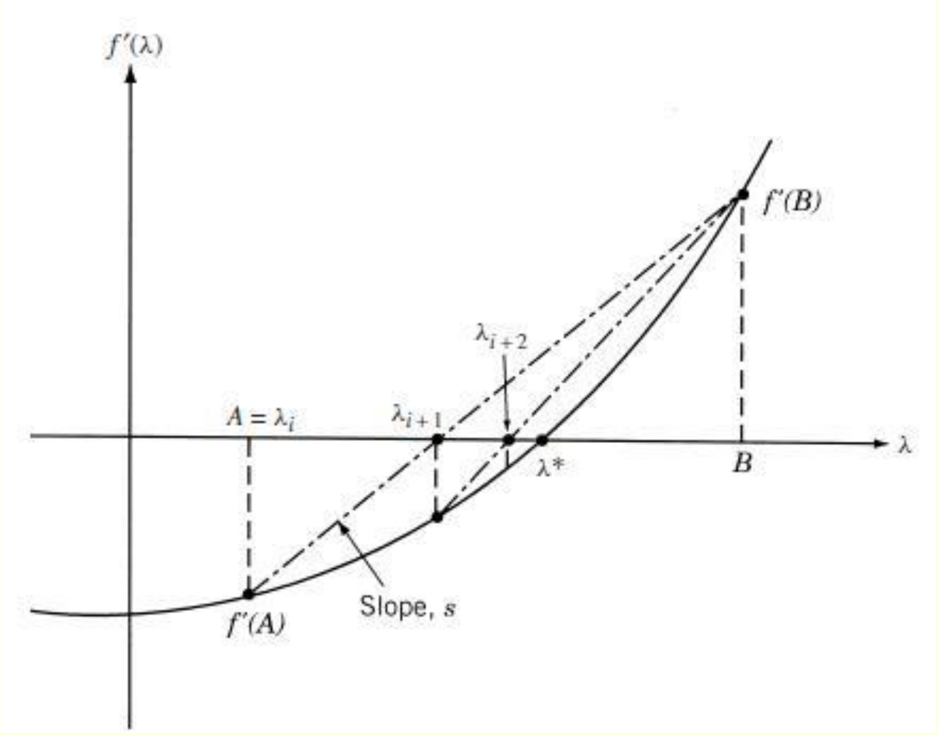
Convergence check:

$$|f'(\lambda_4)| = \left|\frac{f_4^+ - f_4^-}{2\Delta\lambda}\right| = 0.000350 < \varepsilon$$

Since the process has converged, we take the optimum solution as $\lambda^* \approx \lambda_4 = 0.480600$.

Secant

- **Purpose is to bracket the root – so choose 2 points from 2 sides of the root – assume a straight line between them and find the point where the straight line touches the x-axis**



$$\frac{y - f'(A)}{x - A} = \frac{f'(A) - f'(B)}{A - B}$$

$$x = A - \frac{f'(A)(A - B)}{(f'(A) - f'(B))}$$

Secant (contd.)

1. Set $\lambda_1 = A = 0$ and evaluate $f'(A)$. The value of $f'(A)$ will be negative. Assume an initial trial step length t_0. Set $i = 1$.
2. Evaluate $f'(t_0)$.
3. If $f'(t_0) < 0$, set $A = \lambda_i = t_0$, $f'(A) = f'(t_0)$, new $t_0 = 2t_0$, and go to step 2.
4. If $f'(t_0) \geq 0$, set $B = t_0$, $f'(B) = f'(t_0)$, and go to step 5.
5. Find the new approximate solution of the problem as

$$\lambda_{i+1} = A - \frac{f'(A)(B - A)}{f'(B) - f'(A)}$$

6. Test for convergence:

$$|f'(\lambda_i + 1)| \leq \varepsilon$$

where ε is a small quantity. If Eq. (5.75) is satisfied, take $\lambda^* \approx \lambda_{i+1}$ and stop the procedure. Otherwise, go to step 7.

7. If $f'(\lambda_{i+1}) \geq 0$, set new $B = \lambda_{i+1}$, $f'(B) = f'(\lambda_{i+1})$, $i = i + 1$, and go to step 5.
8. If $f'(\lambda_{i+1}) < 0$, set new $A = \lambda_{i+1}$, $f'(A) = f'(\lambda_{i+1})$, $i = i + 1$, and go to step 5.

Secant - Example

Example 5.14 Find the minimum of the function

$$f(\lambda) = 0.65 - \frac{0.75}{1+\lambda^2} - 0.65\lambda \tan^{-1}\frac{1}{\lambda}$$

using the secant method with an initial step size of $t_0 = 0.1$, $\lambda_1 = 0.0$, and $\varepsilon = 0.01$.

SOLUTION $\lambda_1 = A = 0.0$, $t_0 = 0.1$, $f'(A) = -1.02102$, $B = A + t_0 = 0.1$, $f'(B) = -0.744832$. Since $f'(B) < 0$, we set new $A = 0.1$, $f'(A) = -0.744832$, $t_0 = 2(0.1) = 0.2$, $B = \lambda_1 + t_0 = 0.2$, and compute $f'(B) = -0.490343$. Since $f'(B) < 0$, we set new $A = 0.2$, $f'(A) = -0.490343$, $t_0 = 2(0.2) = 0.4$, $B = \lambda_1 + t_0 = 0.4$, and compute $f'(B) = -0.103652$. Since $f'(B) < 0$, we set new $A = 0.4$, $f'(A) = -0.103652$, $t_0 = 2(0.4) = 0.8$, $B = \lambda_1 + t_0 = 0.8$, and compute $f'(B) = +0.180800$. Since $f'(B) > 0$, we proceed to find λ_2.

Iteration 1

Since $A = \lambda_1 = 0.4$, $f'(A) = -0.103652$, $B = 0.8$, $f'(B) = +0.180800$, we compute

$$\lambda_2 = A - \frac{f'(A)(B-A)}{f'(B) - f'(A)} = 0.545757$$

Convergence check: $|f'(\lambda_2)| = |+0.0105789| > \varepsilon$.

Secant - Example

Iteration 2

Since $f'(\lambda_2) = +0.0105789 > 0$, we set new $A = 0.4$, $f'(A) = -0.103652$, $B = \lambda_2 = 0.545757$, $f'(B) = f'(\lambda_2) = +0.0105789$, and compute

$$\lambda_3 = A - \frac{f'(A)(B - A)}{f'(B) - f'(A)} = 0.490632$$

Convergence check: $|f'(\lambda_3)| = |+0.00151235| < \varepsilon$.

Since the process has converged, the optimum solution is given by $\lambda^* \approx \lambda_3 = 0.490632$.