Numerical Methods

(MTH4002)

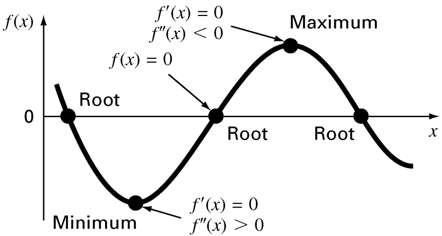
Lecture 06: Numerical Optimization

Dr. Kundan Kumar
Associate Professor
Department of ECE

Faculty of Engineering (ITER)
S'O'A Deemed to be University, Bhubaneswar, India-751030

Introduction

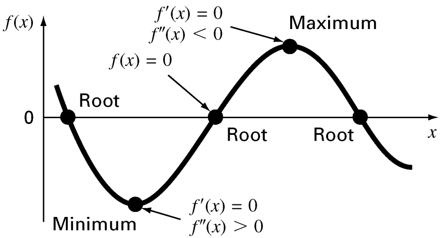
- Root location and optimization are related in the sense that both involve guessing and searching for a point on a function.
- The fundamental difference between the two types of problems is illustrated in Figure.



- Root location involves searching for zeros of a function or functions.
- In contrast, optimization involves searching for either the minimum or the maximum of a function.
- The optimum is the point where the curve is flat.

Introduction

- In mathematical terms, optimum corresponds to the x value where the derivative $f'(x)$ is equal to zero.
- Additionally, the second derivative, $f''(x)$, indicates whether the optimum is a minimum or a maximum:
  - If $f''(x) < 0$, the point is a maximum;
  - If $f''(x) > 0$, the point is a minimum.



- All engineering and science students recall working maxima-minima problems by determining first derivatives of functions in their calculus courses.

Introduction

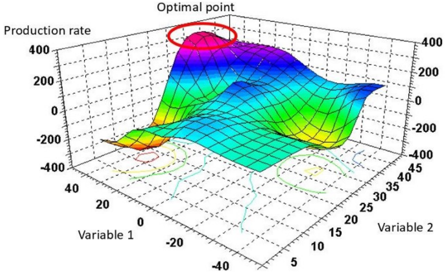
- Engineers must continuously design devices and products that perform tasks in an efficient fashion. In doing so, they are constrained by the limitations of the physical world. Further, they must keep costs down.
- Thus, they are always confronting optimization problems that balance performance and limitations. Some common instances are listed in below.
  - Design aircraft for minimum weight and maximum strength.
  - Optimal trajectories of space vehicles.
  - Design civil engineering structures for minimum cost.
  - Machine learning weight optimization
  - Cost/Error maximization/minimization.

Fundamental elements of optimization problems

- The fundamental elements of optimization problems that will be routinely confronted in engineering practice are

  1. An objective function or cost function
  2. A number of design variables (real number or integers).
  3. Constraints that reflect the limitations that need to be considered while minimizing/maximizing objective/cost function.

maximize	$f(x, y) = xy$
Subject to	$x + 4y = 240$



Some important definitions

- Definition 01: The function f is said to have a local minimum value at $x = p$, if there exists an open interval I containing p so that $f(p) \leq f(x)$ for all $x \in I$. Similarly, f is said to have a local maximum value at $x = p$ if $f(p) \geq f(x)$ for all $x \in I$. If f has either a local minimum or maximum value at $x = p$, it is said to have a local extremum at $x = p$.
- Definition 02: Assume that $f(x)$ is defined on the interval I.
  - (i) If $x_1 < x_2$ implies that $f(x_1) < f(x_2)$ for all $x_1, x_2 \epsilon I$, then f is said to be increasing on I.
  - (ii) If $x_1 < x_2$ implies that $f(x_1) > f(x_2)$ for all $x_1, x_2 \epsilon I$, then f is said to be decreasing on I.

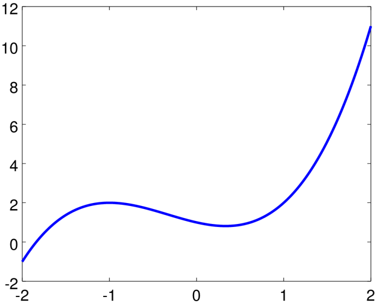
First and Second Derivative Test

- First derivative test: Assume that $f(x)$ is continuous on $I = [a, b]$. Furthermore, suppose that $f'(x)$ is defined for all $x \in (a, b)$, except possibly at $x = p$
  (i) If $f'(x) < 0$ on (a, p) and $f'(x) > 0$ on (p, b), then $f(p)$ is a local minimum.
  (ii) If $f'(x) > 0$ on (a, p) and $f'(x) < 0$ on (p, b), then $f(p)$ is a local maximum.

- Second derivative test: Assume that f is continuous on $[a, b]$ and f' and f'' are defined on (a, b). Also, suppose that $p \in (a, b)$ is a critical point where $f'(p) = 0$.
  (i) If $f''(p) > 0$, then $f(p)$ is a local minimum of f.
  (ii) If $f''(p) < 0$, then $f(p)$ is a local maximum of f.
  (iii) If $f''(p) = 0$, then this test is inconclusive.

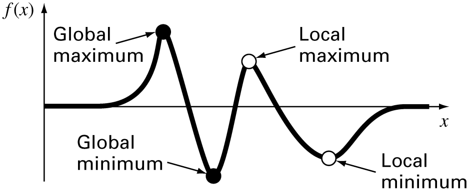
Example

Use the second derivative test to classify the local extrema of $f(x) = x^3 + x^2 - x + 1$ on the interval $[-2, 2]$.



Minimum or maximum of a single variable function

- Recall that root location was complicated by the fact that several roots can occur for a single function.
- Similarly, both local and global optima can occur in optimization. Such cases are called multimodal.
- In almost all instances, we will be interested in finding the absolute highest or lowest value of a function.
- Thus, we must take care that we do not mistake a local result for the global optimum.



Optimization in one dimension

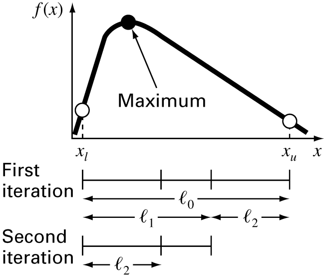
- Single-variable optimization has the goal of finding the value of x that yields an extremum, either a maximum or minimum of $f(x)$.
- Bracketing methods
  - Golden-Ratio Search and Fibonacci Search are examples of a bracketing methods.
  - In these methods, the minimum of $f(x)$ is found out for a given interval by evaluating the function many times and searching for a local minimum.
- Open methods
  - Another method is based on the idea from calculus that the minimum or maximum can be found by solving $f'(x) = 0$.
  - One version of this approach—Newton's method will be discussed.

Golden Ratio Search

- The golden-ratio search is a simple, general-purpose, single-variable search technique.
- For simplicity, we will focus on the problem of finding a maximum.
- When we will discuss the computer algorithm, we will describe the minor modifications needed to simulate a minimum.
- We can start by defining an interval that contains a single answer. That is, the interval should contain a single maximum, and hence is called unimodal.
- We can adopt the nomenclature, where x_l and x_u defined the lower and upper bounds, respectively, of such an interval.

Golden Ratio Search

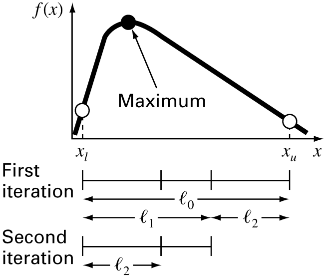
- Rather than using only two function values (which are sufficient to detect a sign change, and hence a zero), we would need third function values to detect whether a maximum occurred.
- Thus, an additional point within the interval has to be chosen.



- Next, we have to pick a fourth point. Then the test for the maximum could be applied to discern whether the maximum occurred within the first three or the last three points.
- The key to making this approach efficient is the wise choice of the intermediate points.

Golden Ratio Search

- This goal can be achieved by specifying that the following two conditions hold.



$$\ell_0 = \ell_1 + \ell_2 \tag{1}$$

$$\frac{\ell_1}{\ell_0} = \frac{\ell_2}{\ell_1} \tag{2}$$

(i) The first condition specifies that the sum of the two sublengths ℓ_1 and ℓ_2 must equal the original interval length.

(ii) The second says that the ratio of the lengths must be equal.

Golden Ratio Search

- Equation (1) can be substituted into Eq. (2), we get

$$\frac{\ell_1}{\ell_1 + \ell_2} = \frac{\ell_2}{\ell_1}$$

- If the reciprocal is taken and $R = \ell_2/\ell_1$, we arrive at

$$1 + R = \frac{1}{R}$$

$$\Rightarrow \qquad R^2 + R - 1 = 0$$

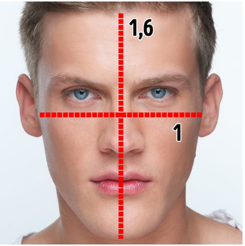
- Which can be solved for the positive root

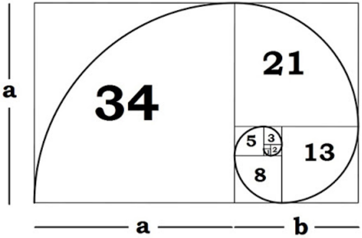
$$R = \frac{-1 + \sqrt{1 - 4(-1)}}{2} = \frac{\sqrt{5} - 1}{2} = 0.61803\ldots$$

- This value, which has been known since antiquity, is called the Golden Ratio.

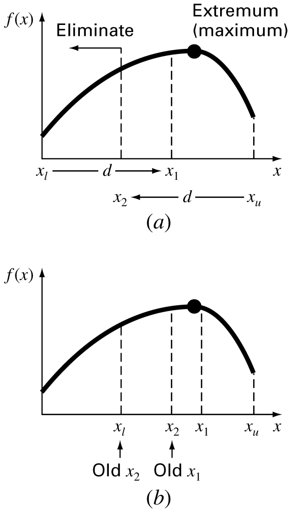
Golden Ratio Search

- Because it allows optima to be found efficiently, it is the key element of the golden-section method we have been developing conceptually.
- Now let us derive an algorithm to implement this approach on the computer.





Golden Ratio Search Algorithm



(1) Start with two initial guesses, x_l and x_u, that bracket one local extremum of $f(x)$. Next, two interior points x_1 and x_2 are chosen according to the golden ratio.

$$\begin{aligned} d &= \frac{\sqrt{5}-1}{2}(x_u - x_l) \\ x_1 &= x_l + d \\ x_2 &= x_u - d \end{aligned} \tag{3}$$

(2) The function is evaluated at these two interior points. Two results can occur:

(i) If $f(x_1) > f(x_2)$ then the domain of x to the left of x_2, from x_l to x_2, can be eliminated because it does not contain the maximum. For this case, x_2 becomes the new x_l for the next iteration.

(ii) If $f(x_2) > f(x_1)$, then the domain of x to the right of x_1, from x_1 to x_u would have been eliminated. In this case, x_1 becomes the new x_u for the next iteration.

Real benefit from the use of the golden ratio

- Because the original x_1 and x_2 were chosen using the golden ratio, we do not have to recalculate all the function values for the next iteration.
- The old x_1 becomes the new x_2. This means that we already have the value for the new $f(x_2)$, since it is the same as the function value at the old x_1.
- To complete the algorithm, we now only need to determine the new x_1. This is done with the same proportionality as before,

$$x_1 = x_l + \frac{\sqrt{5}-1}{2}(x_u - x_l) \tag{4}$$

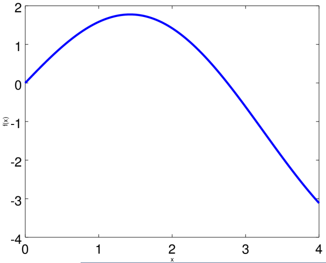
A similar approach would be used for the alternate case where the optimum fell in the left subinterval.

Example

Use the golden-section search to find the maximum of

$$f(x) = 2\sin x - \frac{x^2}{10}$$

within the interval $x_l = 0$ and $x_u = 4$.



Fibonacci Search

- Fibonacci search method differs from the golden ratio method in that the value of r is not constant on each subinterval.
- Additionally, the number of subintervals (iterations)is predetermined and based on the specified tolerances.

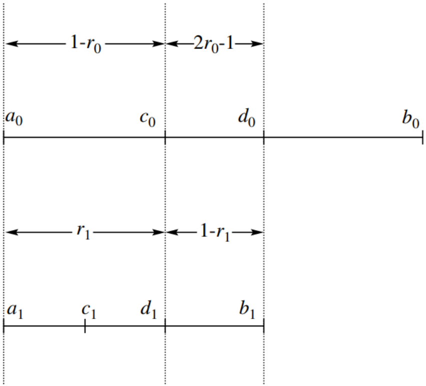
$$\begin{aligned} F_0 = 0, F_1 = 1 \\ F_n = F_{n-1} + F_{n-2} \end{aligned} \tag{5}$$

$n = 2, 3, \ldots$. Thus the Fibonacci numbers are

$$1, 2, 3, 5, 8, 13, 21, \ldots$$

Fibonacci Search

- Assume we are given a function $f(x)$ that is unimodal on the interval $[a_0, b_0]$.
- As in the golden ratio search a value $r_0\,(1/2 < r_0 < 1)$ is selected so that both of the interior points c_0 and d_0 will be used in the next subinterval and there will be only one new function evaluation.
- Without loss of generality assume that $f(c_0) > f(d_0)$. It follows that $a_1 = a_0, b_1 = d_0$, and $d_1 = c_0$.



Fibonacci Search

- If there is to be only one new function evaluation, then we select $r_1\,(1/2 < r_1 < 1)$ for the subinterval $[a_1, b_1]$, such that

$$d_0 - c_0 = b_1 - d_1$$
$$(2r_0 - 1)(b_0 - a_0) = (1 - r_1)(b_1 - a_1)$$
$$(2r_0 - 1)(b_0 - a_0) = (1 - r_1)(r_0(b_0 - a_0))$$
$$2r_0 - 1 = (1 - r_1)\,r_0$$
$$r_1 = \frac{1 - r_0}{r_0}$$

Fibonacci Search

- Substituting $r_0 = F_{n-1}/F_n, n \geq 4$, into this last equation yields

$$r_1 = \frac{1 - \frac{F_{n-1}}{F_n}}{\frac{F_{n-1}}{F_n}} = \frac{F_n - F_{n-1}}{F_{n-1}}$$
$$= \frac{F_{n-2}}{F_{n-1}}$$

- Since, by equation (5), $F_n = F_{n-1} + F_{n-2}$, it follows that the Fibonacci search can be begun with $r_0 = F_{n-1}/F_n$ and continued using $r_k = F_{n-1-k}/F_{n-k}$ for $k = 1, 2, \ldots, n-3$.
- Note that $r_{n=3} = F_2/F_3 = 1/2$, thus no new points can be added at this stage. Therefore, there are a total of $(n-3)+1 = n-2$ steps in this process.

Fibonacci Search

- The $(k+1)$st subinterval is obtained by reducing the length of the kth subinterval by a factor of $r_k = F_{n-1-k}/F_{n-k}$. The length of the last subinterval is

$$\frac{F_{n-1}F_{n-2}\cdots F_2}{F_nF_{n-1}\cdots F_3}(b_0-a_0) = \frac{F_2}{F_n}(b_0-a_0)$$
$$= \frac{1}{F_n}(b_0-a_0) = \frac{b_0-a_0}{F_n}$$

Fibonacci Search

- If the abscissa of the minimum is to be found with a tolerance of ϵ, then we need to find the smallest value of n such that
$$\frac{b_0 - a_0}{F_n} < \epsilon \quad \text{or} \quad F_n > \frac{b_0 - a_0}{\epsilon} \tag{6}$$
The interior points c_k and d_k of the k th subinterval $[a_k, b_k]$ are found, as needed, using the formulas
$$c_k = a_k + \left(1 - \frac{F_{n-k-1}}{F_{n-k}}\right)(b_k - a_k) \tag{7}$$
$$d_k = a_k + \frac{F_{n-k-1}}{F_{n-k}}(b_k - a_k) \tag{8}$$
- Note that the value of n used in formulas (7) and (8) is found using inequality (6).

Fibonacci Search

- Each iteration requires the determination of two new interior points, one from the previous iteration and the second from formula (7) and (8). When $r_0 = F_2/F_3 = 1/2$ the two interior points will be concurrent in the middle of the interval. To distinguish the two interior points a small distinguishability constant, e, is introducted. Thus when formula (7) and (8) is used, the coefficients of $(b_k - a_k)$ are $1/2 - e$ or $1/2 + e$, respectively.

Example

Find the minimum of the function $f(x) = x^2 - \sin(x)$ on the interval [0,1] using the Fibonacci search method. Use a tolerance of $\epsilon = 10^{-4}$ and the distinguishability constant $e = 0.01$

Solution:

The smallest Fibonacci number satisfying

$$F_n > \frac{b_0 - a_0}{\epsilon} = \frac{1-0}{10^{-4}} = 10,000$$

is $F_{21} = 10,946$. Thus $n = 21$.

Exmple

Given $a_0 = 0$ and $b_0 = 1$. Formulas (7) and (8) yield

$$c_0 = 0 + \left(1 - \frac{F_{20}}{F_{21}}\right)(1-0) \approx 0.3819660$$
$$d_0 = 0 + \frac{F_{20}}{F_{21}}(1-0) \approx 0.6180340$$

Then set $a_1 = a_0, b_1 = d_0$, and $d_1 = c_0$, since $f(0.3819660) = -0.2268475$ and $f(0.6180340) = -0.1974679\,(f(d_0) \geq f(c_0))$.

The new subinterval containing the abscissa of the minimum of f is $[a_1, b_1] = [0, 0.6180340]$. Now use formula (7) to calculate the interior point c_1 :

$$c_1 = a_1 + \left(1 - \frac{F_{21-1-1}}{F_{21-1}}\right)(b_1 - a_1)$$
$$= 0 + \left(1 - \frac{F_{19}}{F_{20}}\right)(0.6180340 - 0) \approx 0.2360680$$

Example

Now compute and compare $f(c_1)$ and $f(d_1)$ to determine the new subinterval $\lfloor a_2, b_2 \rfloor$, and continue the iteration process. Some of the computations are shown in the Table.

k	a_k	c_k	d_k	b_k
0	0.0000000	0.3819660	0.6180340	1.0000000
1	0.0000000	0.2360680	0.3819660	0.6180340
2	0.2360680	0.3819660	0.4721359	0.6180340
3	0.3819660	0.4721359	0.5278641	0.6180340
4	0.3819660	0.4376941	0.4721359	0.5278641
⋮	⋮	⋮	⋮	⋮
16	0.4499360	0.4501188	0.4502102	0.4503928
17	0.4501188	0.4502101	0.4503015	0.4503928
18	0.4501188	0.4502083	0.4502101	0.4503015

Example

At the seventeenth iteration the interval has been narrowed down to $[a_{17}, b_{17}] = [0.4501188, 0.4503928]$, where $c_{17} = 0.4502101, d_{17} = 0.4503105$, and $f(d_{17}) \geq f(c_{17})$ Thus $[a_{18}, b_{18}] = [0.4501188, 0.4503015]$ and $d_{18} = 0.4502101$. At this stage the multiplier is $r_{18} = 1 - F_2/F_3 = 1 - 1/2 = 1/2$ and the distinguishability constant $e = 0.01$ is used to calculate c_{18} :

$$\begin{aligned} c_{18} &= a_{18} + (0.5 - 0.01)(b_{18} - a_{18}) \\ &= 0.4501188 - 0.49(0.450315 - 0.4501188) \\ &\approx 0.4502083 \end{aligned}$$

since $f(d_{18}) \geq f(c_{18})$, the final subinterval is $[a_{19}, b_{19}] = [0.4501188, 0.4502101]$. This interval has width 0.0000913 . We choose to report the abscissa of the minimum as the midpoint of this interval. Therefore, the minimum value is $f(0.4501645) = -0.2324656$

Example

Remarks

- Both the Fibonacci and golden ratio search methods can be applied in cases where $f(x)$ is not differentiable. It should be noted that when n is small the Fibonacci method is more efficient than the golden ratio method. However, for n large the two methods are almost identical.

Gradient and Newton's Methods

- Recall that the Newton-Raphson method of is an open method that finds the root x of a function such that $f(x) = 0$. The method is summarized as

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$$

A similar open approach can be used to find an optimum of $f(x)$ by defining a new function, $g(x) = f'(x)$. Thus, because the same optimal value x^* satisfies both

$$f'(x^*) = g(x^*) = 0$$

we can use the following,

$$x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)}$$

as a technique to find the minimum or maximum of $f(x)$.

Gradient and Newton's Methods

- Newton's method is an open method similar to Newton-Raphson because it does not require initial guesses that bracket the optimum.
- In addition, it also shares the disadvantage that it may be divergent.
- Finally, it is usually a good idea to check that the second derivative has the correct sign to confirm that the technique is converging on the result you desire.

Example

Use Newton's method to find the maximum of

$$f(x) = 2\sin x - \frac{x^2}{10}$$

with an initial guess of $x_0 = 2.5$.

Solution

The first and second derivatives of the function can be evaluated as

$$f'(x) = 2\cos x - \tfrac{x}{5}$$
$$f''(x) = -2\sin x - \tfrac{1}{5}$$

Example

which can be substituted into governing equation to give

$$x_{i+1} = x_i - \frac{2\cos x_i - (x_i/5)}{-2\sin x_i - (1/5)}$$

Substituting the initial guess yields

$$x_1 = 2.5 - \frac{2\cos(2.5) - (2.5/5)}{-2\sin(2.5) - (1/5)} = 0.99508$$

which has a function value $f(x_1)$ as 1.57859.

Example

The second iteration gives

$$x_1 = 0.995 - \frac{2\cos(0.995) - (0.995/5)}{-2\sin(0.995) - (1/5)} = 1.46901$$

which has a function value $f(x_2)$ 1.77385 . The process can be repeated, with the results tabulated below:

i	x	f(x)	f′(x)	f″(x)
0	2.5	0.57194	−2.10229	−1.39694
1	0.99508	1.57859	0.88985	−1.87761
2	1.46901	1.77385	−0.09058	−2.18965
3	1.42764	1.77573	−0.00020	−2.17954
4	1.42755	1.77573	0.00000	−2.17952

References

- Numerical Methods Using MATLAB by Matthews and Fink, Pearson
- Applied Numerical Methods with MATLAB for Engineers and Scientists, Third Edition Steven C. Chapra, McGraw-Hill
- Numerical Methods for Engineers and Scientists An Introduction with Applications using MATLAB, Third Edition, Amos Gilat and Vish Subramaniam, John Wiley & Sons

Thank you!