Chong, Edwin Kah Pin., and Stanislaw H. Żak. *An Introduction to Optimization*. Fourth edition., John Wiley & Sons, Inc., 2013.

2 ONE-DIMENSIONAL SEARCH METHODS

2.1 GOLDEN SECTION SEARCH

- The search methods allow us to determine the minimizer of a function $f: \mathbb{R} \to \mathbb{R}$ over a closed interval, say $[a_0, b_0]$.

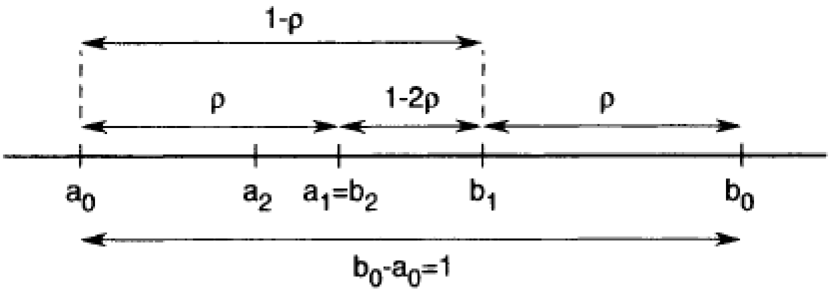
- We assume that the objective function f is unimodal, which means that f has only one local minimizer.

- We choose the intermediate points in such a way that the reduction in the range is symmetric, in the sense that

$$a_1 - a_0 = b_0 - b_1 = \rho(b_0 - a_0),$$

where $\rho < 1/2$.

- If $f(a_1) < f(b_1)$, then the minimizer must lie in the range $[a_0, b_1]$. If, on the other hand, $f(a_1) \geq f(b_1)$, then the minimizer is located in the range $[a_1, b_0]$.



- Without loss of generality, imagine that the original range $[a_0, b_0]$ is of unit length. Then, to have only one new evaluation of f, it is enough to choose ρ so that

$$\rho(b_1 - a_0) = b_1 - b_2.$$

- Because $b_1 - a_0 = 1 - \rho$ and $b_1 - b_2 = 1 - 2\rho$, we have

$$\begin{aligned} \rho(b_1 - a_0) &= b_1 - b_2 \\ \rho(1 - \rho) &= 1 - 2\rho \\ \rho^2 - 3\rho + 1 &= 0 \\ \rho &= \frac{3 \pm \sqrt{5}}{2} \end{aligned}$$

- Because we require that $\rho < 1/2$, we take $\rho = \left(3 - \sqrt{5}\right)/2 \approx 0.382$.

- Using the golden section rule means that at every stage of the uncertainty range reduction (except the first), the objective function/need only be evaluated at one new point. The uncertainty range is reduced by the ratio $1 - \rho \approx 0.61803$ at every stage. Hence, N steps of reduction using the golden section method reduces the range by the factor $(1 - \rho)^N \approx (0.61803)^N$.

Chong, Edwin Kah Pin., and Stanislaw H. Żak. *An Introduction to Optimization*. Fourth edition., John Wiley & Sons, Inc., 2013.

Example 2.1.1. Use the golden section search to find the value of x that minimizes

$$f(x) = x^4 - 14x^3 + 60x^2 - 70x$$

in the range $[0, 2]$. Locate this value of x to within a range of 0.3.

After N stages the range $[0, 2]$ is reduced by $(0.61803)^N$. So, we choose N so that

$$(0.61803)^N \leq 0.3/2.$$

Four stages of reduction will do; that is, $N = 4$.

Iteration 1.

$$f(a_1) = -24.36,$$
$$f(b_1) = -18.96,$$

where $a_1 = a_0 + \rho(b_0 - a_0) = 0.7639$ and $b_1 = a_0 + (1-\rho)(b_0 - a_0) = 1.236$. Thus, $f(a_1) < f(b_1)$, so the uncertainty interval is reduced to $[a_0, b_1] = [0, 1.236]$.

Iteration 2.

$$f(a_2) = -21.10,$$
$$f(b_2) = f(a_1) = -24.36,$$

where $a_2 = a_0 + \rho(b_1 - a_0) = 0.4721$. Thus, $f(a_2) > f(b_2)$, so the uncertainty interval is reduced to $[a_2, b_1] = [0.4721, 1.236]$.

Iteration 3.

$$f(a_3) = f(b_2) = -24.36,$$
$$f(b_3) = -23.59,$$

where $b_3 = a_2 + (1-\rho)(b_1 - a_2) = 0.9443$. Thus, $f(a_3) < f(b_3)$, so the uncertainty interval is reduced to $[a_2, b_3] = [0.4721, 0.9443]$.

Iteration 4.

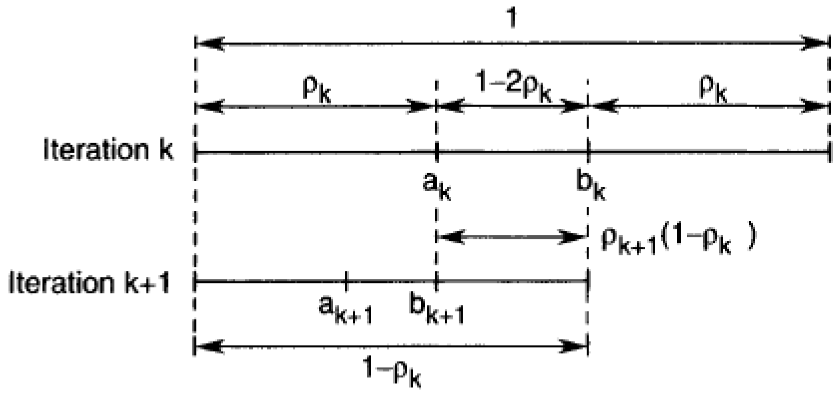
$$f(a_4) = -23.84,$$
$$f(b_4) = f(a_3) = -24.36,$$

where $a_4 = a_2 + \rho(b_3 - a_2) = 0.6525$. Thus, $f(a_4) > f(b_4)$, so the uncertainty interval is reduced to $[a_4, b_3] = [0.6525, 0.9443]$.

Note that $b_3 - a_4 = 0.292 < 0.3$.

2.2 FIBONACCI SEARCH

- Suppose now that we are allowed to vary the value ρ from stage to stage, so that at the kth stage in the reduction process we use a value ρ_k, at the next stage we use a value ρ_{k+1}, and so on.



- Without loss of generality, imagine that the range $[a_{k-1}, b_{k-1}]$ is of unit length. Then, to have only one new evaluation of f, it is enough to choose ρ_{k+1} so that

$$\rho_{k+1}(b_k - a_{k-1}) = b_k - b_{k+1}.$$

- Because $b_k - a_{k-1} = 1 - \rho_k$ and $b_k - b_{k+1} = 1 - 2\rho_k$, we have

$$\begin{aligned} \rho_{k+1}(b_k - a_{k-1}) &= b_k - b_{k+1} \\ \rho_{k+1}(1-\rho_k) &= 1 - 2\rho_k \\ \rho_{k+1} &= \frac{1-\rho_k-\rho_k}{1-\rho_k} \\ \rho_{k+1} &= 1 - \frac{\rho_k}{1-\rho_k} \end{aligned}$$

- Then, after N iterations of the algorithm, the uncertainty range is reduced by a factor of $(1-\rho_1)(1-\rho_2)\cdots(1-\rho_N)$.

- The sequence $\rho_1, \rho_2, \ldots$ that minimizes the factor above is

$$\rho_1 = 1 - \frac{F_N}{F_{N+1}}, \rho_2 = 1 - \frac{F_{N-1}}{F_N}, \ldots, \rho_k = 1 - \frac{F_{N-k+1}}{F_{N-k+2}}, \ldots, \rho_N = 1 - \frac{F_1}{F_2},$$

  where the F_k are the elements of the Fibonacci sequence.

- The **Fibonacci sequence** is defined as follows. First, let $F_{-1} = 0$ and $F_0 = 1$ by convention. Then, for $k \geq 0$, $F_{k+1} = F_k + F_{k-1}$.

- Some values of elements in the Fibonacci sequence are as follows: $F_1 = 1, F_2 = 2, F_3 = 3, F_4 = 5, F_5 = 8, F_6 = 13, F_7 = 21, F_8 = 34$.

- In the Fibonacci search method, the uncertainty range is reduced by the factor:

$$(1-\rho_1)(1-\rho_2)\cdots(1-\rho_N) = \frac{F_N}{F_{N+1}}\frac{F_{N-1}}{F_N}\cdots\frac{F_1}{F_2} = \frac{1}{F_{N+1}}.$$

- Fibonacci method is better than the golden section method in that it gives a smaller final uncertainty range.

- In the final iteration of the Fibonacci search method, because $\rho_N = 1 - F_1/F_2 = 1/2$, the two intermediate points coincide in the middle of the uncertainty interval, and therefore we cannot further reduce the uncertainty range.

- To get around this problem, we perform the new evaluation for the last iteration using $\rho_N = 1/2 - \varepsilon$, where ε is a small number. Therefore, in the worst case, the reduction factor in the uncertainty range for the Fibonacci method is

$$(1-\rho_1)(1-\rho_2)\cdots(1-\rho_N) = \frac{F_N}{F_{N+1}}\frac{F_{N-1}}{F_N}\cdots\frac{F_2}{F_3}\left(\frac{F_1}{F_2}+\varepsilon\right) = \frac{1+2\varepsilon}{F_{N+1}}.$$

Example 2.1.1. Use the Fibonacci search to find the value of x that minimizes

$$f(x) = x^4 - 14x^3 + 60x^2 - 70x$$

in the range $[0, 2]$. Locate this value of x to within a range of 0.3.

After N stages the range $[0, 2]$ is reduced by $(1+2\varepsilon)/F_{N+1}$. So, we choose N so that

$$\frac{1+2\varepsilon}{F_{N+1}} \le \frac{0.3}{2}.$$

Let $\varepsilon = 0.1$. Four stages of reduction will do; that is, $N = 4$.

Iteration 1.

$$f(a_1) = \cancel{-24.36} - 24.340,$$
$$f(b_1) = \cancel{-18.96} - 18.652,$$

where $a_1 = a_0 + \rho(b_0 - a_0) = \cancel{0.7639}0.750$ and $b_1 = a_0 + (1-\rho)(b_0 - a_0) = \cancel{1.236}\, 1.250$. Thus, $f(a_1) < f(b_1)$, so the uncertainty interval is reduced to $[a_0, b_1] = [0, \cancel{1.236}1.250]$.

Iteration 2.

$$f(a_2) = \cancel{-21.10} - 21.688,$$
$$f(b_2) = f(a_1) = \cancel{-24.36} - 24.340,$$

where $a_2 = a_0 + \rho(b_1 - a_0) = \cancel{0.4721}0.500$. Thus, $f(a_2) > f(b_2)$, so the uncertainty interval is reduced to $[a_2, b_1] = [\cancel{0.4721}0.500, \cancel{1.236}1.250]$.

Iteration 3.

$$f(a_3) = f(b_2) = \cancel{-24.36} - 24.340,$$
$$f(b_3) = \cancel{-23.59} - 23.000,$$

where $b_3 = a_2 + (1-\rho)(b_1 - a_2) = \cancel{0.9443}1.000$. Thus, $f(a_3) < f(b_3)$, so the uncertainty interval is reduced to $[a_2, b_3] = [\cancel{0.4721}0.500, \cancel{0.9443}1.000]$.

Iteration 4.

$$f(a_4) = \cancel{-23.84} - 24.162,$$
$$f(b_4) = f(a_3) = \cancel{-24.36} - 24.340,$$

where $a_4 = a_2 + \rho(b_3 - a_2) = \cancel{0.65250}0.700$. Thus, $f(a_4) > f(b_4)$, so the uncertainty interval is reduced to $[a_4, b_3] = [\cancel{0.65250}0.700, \cancel{0.94431}1.000]$.

Note that $b_3 - a_4 = 0.3 \leq 0.3$.

2.3 NEWTON'S METHOD

- Assume that at each measurement point $x^{(k)}$ we can calculate $f(x^{(k)})$, $f'(x^{(k)})$ and $f''(x^{(k)})$. We can fit a quadratic function through $x^{(k)}$ that matches its first and second derivatives with that of the function f. This quadratic has the form

$$q(x) = f(x^{(k)}) + f'(x^{(k)})(x - x^{(k)}) + \frac{1}{2} f''(x^{(k)})(x - x^{(k)})^2.$$

- Note that $q(x^{(k)}) = f(x^{(k)})$, $q'(x^{(k)}) = f'(x^{(k)})$, and $q''(x^{(k)}) = f''(x^{(k)})$. Then, instead of minimizing f, we minimize its approximation q.

- The first-order necessary condition for a minimizer of q yields

$$0 = q'(x) = f'(x^{(k)}) + f''(x^{(k)})(x - x^{(k)}).$$

- Setting $x = x^{(k+1)}$, we obtain

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

- Newton's method works well if $f''(x) > 0$ everywhere. However, if $f''(x) < 0$ for some x, Newton's method may fail to converge to the minimizer.

Example 2.3.1. Use Newton's method, find the minimizer of

$$f(x) = \frac{1}{2}x^2 - \sin x.$$

The initial value is $x^{(0)} = 0.5$. The required accuracy is $\epsilon = 10^{-5}$, in the sense that we stop when $|x^{(k+1)} - x^{(k)}| < \epsilon$.

We compute

$$f'(x) = x - \cos x, \quad f''(x) = 1 + \sin x.$$

Hence,

$$x^{(1)} = x^{(0)} - \frac{f'(x^{(0)})}{f''(x^{(0)})} = 0.5 - \frac{0.5 - \cos 0.5}{1 + \sin 0.5} = 0.7552$$

$$x^{(2)} = x^{(1)} - \frac{f'(x^{(1)})}{f''(x^{(1)})} = x^{(1)} - \frac{0.02710}{1.685} = 0.7391$$

$$x^{(3)} = x^{(2)} - \frac{f'\left(x^{(2)}\right)}{f''\left(x^{(2)}\right)} = x^{(2)} - \frac{9.461 \times 10^{-5}}{1.673} = 0.7390$$

$$x^{(4)} = x^{(3)} - \frac{f'\left(x^{(3)}\right)}{f''\left(x^{(3)}\right)} = x^{(3)} - \frac{1.17 \times 10^{-9}}{1.673} = 0.7390$$

Note that $\left|x^{(4)} - x^{(3)}\right| < 10^{-5}$. Furthermore, $f'\left(x^{(4)}\right) = -8.6 \times 10^{-6} \approx 0$. Observe that $f''\left(x^{(4)}\right) = 1.673 > 0$, so we can assume that $x^* \approx x^{(4)}$ is a strict minimizer.

- Newton's method can also be viewed as a way to drive the first derivative of f to zero. Indeed, if we set $g(x) = f'(x)$, then we obtain a formula for iterative solution of the equation $g(x) = 0$:

$$x^{(k+1)} = x^{(k)} - \frac{g\left(x^{(k)}\right)}{g'\left(x^{(k)}\right)}.$$

Example 2.3.2. Apply Newton's method to improve a first approximation, $x^{(0)} = 12$, to the root of the equation

$$g(x) = x^3 - 12.2x^2 + 7.45x + 42 = 0.$$

We have $g'(x) = 3x^2 - 24.4x + 7.45$.

$$x^{(1)} = x^{(0)} - \frac{g\left(x^{(0)}\right)}{g'\left(x^{(0)}\right)} = 12 - \frac{102.6}{146.65} = 11.33$$

$$x^{(2)} = x^{(1)} - \frac{g\left(x^{(1)}\right)}{g'\left(x^{(1)}\right)} = 11.33 - \frac{14.73}{116.11} = 11.21.$$

2.4 SECANT METHOD

- Newton's method for minimizing f uses **second derivatives** of f.
- If the second derivative is not available, we may attempt to approximate it using first derivative information. In particular, we may approximate $f''\left(x^{(k)}\right)$ with

$$\frac{f'\left(x^{(k)}\right) - f'\left(x^{(k-1)}\right)}{x^{(k)} - x^{(k-1)}}.$$

- Using the foregoing approximation of the second derivative, we obtain the algorithm

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

called the **secant method**.

- Note that the algorithm requires two initial points $x^{(-1)}$ and $x^{(0)}$.

Example 2.4.1. Apply the secant method to find the root of the equation

$$g(x) = x^3 - 12.2x^2 + 7.45x + 42 = 0,$$

with starting points $x^{(-1)} = 13$ and $x^{(0)} = 12$.

We have $g'(x) = 3x^2 - 24.4x + 7.45$.

$$x^{(1)} = 11.40,$$

$$x^{(2)} = 11.25.$$

2.5 REMARKS ON LINE SEARCH METHODS

- One-dimensional search methods play an important role in multidimensional optimization problems. Iterative algorithms for solving such optimization problems typically involve a line search at every iteration.

- Let $f: \mathbb{R}^n \to \mathbb{R}$ be a function that we wish to minimize. Iterative algorithms for finding a minimizer of f are of the form

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k+1)} + \alpha_k \boldsymbol{d}^{(k)},$$

  where $\boldsymbol{x}^{(0)}$ is a given initial point and $\alpha_k \geq 0$ is chosen to minimize $\phi_k(\alpha) = f\left(\boldsymbol{x}^{(k)} + \alpha \boldsymbol{d}^{(k)}\right)$. The vector $\boldsymbol{d}^{(k)}$ is called the **search direction**.

- Note that choice of α_k involves a one-dimensional minimization.